

Mentale Verursachung

Problemaufriss und Darstellung einer Debatte innerhalb der zeitgenössischen
Philosophie des Geistes

Inauguraldissertation
zur Erlangung des Doktorgrades der Philosophie
an der Ludwig-Maximilians-Universität München

vorgelegt von
Stefan Hammling
aus Oettingen i. Bay.

2018

Erstgutachter: Prof. Dr. Stephan Sellmaier

Zweitgutachter: Prof. Dr. Axel Hutter

Datum der mündlichen Prüfung: 20.07.2015

Inhaltsverzeichnis

<u>Einleitung</u>	5
Was ist Kausalität?	6
Kausalitätstheorien	11
Zielsetzung der Arbeit und ausführliche Darstellung der Vorgehensweise	20
<u>Das Interaktionsproblem</u>	40
Descartes' Theorie von Körper und Geist	40
Facetten des Interaktionsproblems	42
Räumliche Lokalisierbarkeit und kausaler Nexus	43
Impuls- bzw. Energieerhaltung und Vollständigkeit der Physik	44
Descartes' Erklärung psychophysischer Interaktion innerhalb seiner Theorie	46
Weitere Lösungen des Interaktionsproblems aus Sicht des Substanzdualismus	49
Das Pairing Problem	52
Kritik am und Lösungen für das Pairing Problem	58
Schwache vs. starke Nicht-Intelligibilität	58
Dispositionale Eigenschaften als Lösung des Pairing Problems	59
Seelenfeld oder Mentalraum	62
Individualistische Kräfte	66
Die nomologische Annahme und personengebundene Gesetze	70
<u>Das Problem um die Anomalie des Mentalen</u>	78
Davidsons anomaler Monismus als Lösung für das Problem um die Anomalie des Mentalen	80
Ist das Mentale qua mentalem kausal relevant?	92
Kontrafaktische Abhängigkeit als Lösung für das Anomalie-Problem	99
Sind Kausalgesetze notwendigerweise strikte Gesetze?	111
<u>Das Exklusionsproblem</u>	127
Kritik am Exklusionsargument	152
Überdeterminierung I	152
Versickern die kausalen Kräfte?	160
Überdeterminierung II	172
Mögliche Lösungsstrategien für das Exklusionsproblem	180
Determinablen und Determinaten	180
Kausale Wirksamkeit und kausale Relevanz	188
Tropen-Monismus	202
Mentale Quausalität	212
<u>Externalismus und Anti-Individualismus als Problem für mentale Verursachung</u>	225
Exkurs: Externalismus und Anti-Individualismus	225
Externalismus – Hilary Putnam	225
Anti-Individualismus – Tyler Burge	231
Externalismus und Anti-Individualismus als Problem für mentale Verursachung	237
Lösungen für das Externalismus/Anti-Individualismus-Problem	242
Enger Gehalt mentaler Zustände	242
Erklärung vs. Verursachung	253

SCHLUSS	264
AUSLÖSENDE UND STRUKTURIERENDE URSACHEN	264
LITERATURVERZEICHNIS	281

Einleitung

Es hat den Anschein als wäre mentale Verursachung¹ eine unumstößliche Tatsache der Erlebenswelt jedes² Menschen. So erfahren wir beispielsweise Wünsche als Ursachen für Körperbewegungen, Verletzungen als Ursache von Schmerzen, oder, Gedanken als Ursache von anderen Gedanken. Da der Zugang zu diesem inneren Erleben direkt ist, scheint die Möglichkeit eines Irrtums darüber ausgeschlossen zu sein. Während man etwa bei der Einschätzung der Art einer Verletzung durchaus falsch liegen mag, wäre die Behauptung man könne sich darüber täuschen, dass man Schmerzen empfindet, zunächst kontrain intuitiv. Schließlich ist das Hauptindiz für die Existenz von Schmerzen, dass man sie fühlt oder eben nicht. In anderen Worten: wer *glaubt* Schmerzen zu haben, *hat* de facto Schmerzen. Ähnlich verhält es sich auch mit Wünschen, Absichten, Empfindungen, etc.

Dieser privilegierte Zugang zum eigenen Bewusstsein ist beispielsweise auch Descartes' Ausgangspunkt für seine berühmte Schlussfolgerung Cogito ergo sum. Dem Argument zufolge ist das Haben von Bewusstsein die einzige Tatsache, die über jeden skeptischen Zweifel erhaben sein kann und die somit auch den Ausgangspunkt für die Philosophie generell darstellt. Schließlich verkörpert das Bewusstsein auch den Wesenskern der Persönlichkeit, denn das Selbst konstituiert sich im Wesentlichen aus der Innenwahrnehmung des Menschen als eine denkende,ühlende und Handlungen verursachende Einheit. Das Bindeglied zwischen dem Selbst und der Welt ist offensichtlich der Körper. Mentale Verursachung wiederum gewährleistet die Verquickung von Selbst und Körper. So könnte eine – zugegebenermaßen stark vereinfachte und verkürzte Darstellung – von Descartes' Gedankengang aussehen. Historisch betrachtet war Descartes der erste, der die Frage nach dem Verhältnis von Seele und Körper in dieser Form aufgeworfen hat. Derzeit würde man diesen Themenkomplex als ‚Leib-Seele-Problem‘ oder als ‚Körper-Geist-Problem‘ bezeichnen. In dessen Gravitationszentrum steht die Frage nach der Möglichkeit mentaler Verursachung.

¹ Zunächst verstehe ich unter dem Begriff ‚mentale Verursachung‘ schlichtweg die Möglichkeit einer kausalen Wechselwirkung zwischen Mentalem und Physikalischem. Natürlich stellt sich sofort die Frage nach einer weiteren Klärung der Begrifflichkeiten – was bedeuten ‚mental‘, ‚physikalisch‘, ‚Kausalität‘ in diesem Zusammenhang? Zum Konzept von Kausalität werde ich im Laufe der Einleitung noch weiter Stellung nehmen. Hinsichtlich der Bedeutung der Begriffe ‚mental‘ und ‚physikalisch‘ bitte ich sich an dem ihnen im alltäglichen Gebrauch zugeschriebenen Sinn zu orientieren.

² Zumindest gilt dies für psychisch gesunde Menschen aus dem abendländisch-westlichen Kulturkreis. In anderen sozio-kulturellen Kontexten ist das Konzept eines handlungsverursachenden Selbst nicht so tief verwurzelt. Dementsprechend mag das Interesse an der Fragestellung nach der Möglichkeit mentaler Verursachung in Gesellschaften, die das Individuum weniger betonen als die okzidentale, auch geringer sein.

Kausale Erklärungen, die die Existenz mentaler Verursachung voraussetzen, sind im Alltag ubiquitär. Hierzu einige Beispiele: ‚Egon griff nach dem Glas, weil er Durst hatte.‘, ‚Um ihre Kopfschmerzen zu lindern, nahm Melissa eine Tablette.‘, ‚Der Geruch von Gegrilltem machte Hans hungrig.‘, ‚Die Gedanken an den Tod ihrer Katze ließen Jolanda traurig werden.‘, ‚Marlenes Selbstbild als zuverlässige Studentin veranlasste sie pünktlich zur Vorlesung aufzubrechen.‘ Diese Liste ließe sich endlos fortsetzen, denn jegliche *kausale* Erklärung einer *Handlung* scheint die Möglichkeit mentaler Kausalität bereits vorauszusetzen. Es ist wohl keine Untertreibung zu behaupten, dass für das Gros zwischenmenschlicher Aktivitäten und Interaktionen kausale Erklärungen eine entscheidende Rolle spielen.

Was aus der Innenperspektive betrachtet als selbstverständlich erscheint, ist bei einer genaueren Betrachtung und philosophischen Analyse jedoch durchaus mit Problemen behaftet. Insbesondere scheint es schwierig zu sein, dem Mentalen einerseits eine gewisse Autonomie zuzuschreiben und es andererseits mit dem Primat der Physik in Einklang zu bringen. Denn mit dem Siegeszug der empirischen Naturwissenschaften wurde ein Weltbild etabliert, in dem es scheinbar keinen Platz mehr gibt für etwas, das nicht vollständig auf Physikalisches reduzierbar ist.

Was ist Kausalität?

Die Suche nach einer Antwort auf die Frage nach dem Wesen von Kausalität ist Stathis Psillos zufolge ein genuin philosophisches Unterfangen und hängt eng mit den Konzepten von ‚Erklärung‘ und ‚Naturgesetz‘ zusammen. Psillos nähert sich dem Thema Kausalität zunächst mit einer Auflistung von Allgemeinplätzen an, derer jedwede Theorie von Verursachung seiner Meinung nach gerecht werden müsse. Anschließend formuliert er zwei fundamentale Intuitionen im Hinblick auf Kausalität. Schließlich zeigt er noch drei Dimensionen auf, anhand derer man die Diskussion über Verursachungstheorien strukturieren kann.

Psillos’ Auflistung der Allgemeinplätze umfasst folgende Punkte:

- The *difference* platitude: causes make a difference – that is, things would be different if the causes of some effects were absent. This platitude is normally cast in two ways: the *counterfactual* way – if the cause hadn’t been, the effect wouldn’t have been either; and the *probabilistic* way – causes raise the *chances* of their effects – that is, the probability that a certain event happens is higher if we take into account its cause than if we don’t.

- The *recipe* platitude: causes are recipes for producing or preventing their effects – that is, causes are the means to produce (or prevent) certain ends (effects). This platitude is normally cast in terms of *manipulability*: causes can be manipulated to bring about certain effects.
- The *explanation* platitude: causes explain their effects, but not vice versa.
- The *evidence* platitude: causes are evidence for their effects – that is, knowing that *c* causes *e*, and knowing that *c* occurred, gives us (some) reason to expect that *e* will occur.³

Die erste triviale Wahrheit im Hinblick auf Kausalität bezeichnet Psillos als den Unterschieds-Allgemeinplatz. Demzufolge muss eine Ursache immer einen Unterschied machen. D.h. irgendetwas im gesamten Weltgeschehen (außer der Tatsache, dass die entsprechende Ursache nicht eingetreten ist) müsste, falls eine Ursache nicht stattfindet, in der Folge divergieren. Diese Erkenntnis kann laut Psillos entweder kontrafaktisch oder probabilistisch zum Ausdruck gebracht werden. Bedient man sich hierzu der Kontrafaktizität, resultiert dies in der Vorstellung, dass, wenn sich die Ursache nicht ereignet hätte, auch die Wirkung ausgeblieben wäre. Nimmt man hingegen Rekurs auf Probabilität, so lautet die Kernaussage, dass Ursachen die Wahrscheinlichkeit von Wirkungen erhöhen. Berücksichtigt man also die Ursache bei der Kalkulation der Wahrscheinlichkeit eines (Wirkungs)Ereignisses, dann ist das Eintreten dieses Ereignisses wahrscheinlicher wenn die Ursache stattgefunden hat, als wenn die Ursache nicht eingetreten wäre. Der zweite Gemeinplatz im Hinblick auf Kausalität ist laut Psillos die sogenannte Rezept-Platitüde. Diese besagt, dass Ursachen sozusagen ‚Gebrauchsanleitungen‘ oder ‚Rezepte‘ für das Bewirken bzw. für das Unterbinden von Effekten sind. Dieser Allgemeinplatz schlägt sich für gewöhnlich in der Forderung nach Manipulierbarkeit nieder, gemäß derer Ursachen in einer Art und Weise kontrolliert werden können, dass sie bestimmte Wirkungen herbeiführen. Eine dritte banale Minimalbedingung für jegliche Kausalitätstheorie besteht Psillos zufolge darin, dass Ursachen ihre Wirkungen erklären müssen, wobei dies umgekehrt nicht der Fall sein muss. Der vierte und letzte Allgemeinplatz, dem laut Psillos jede Kausalitätstheorie genügen muss, ist die Beweis-Platitüde. Diese bekundet, dass Ursachen gewissermaßen Beweise für ihre Wirkungen sind. Ist nämlich bekannt, dass ein Ereignis *c* ein Ereignis *e* verursacht und weiß man zudem, dass *c* eingetreten ist, dann ist es berechtigt anzunehmen, dass auch *e* stattfinden wird. Diese vier soeben beschriebenen Allgemeinplätze sind Psillos zufolge universell gültig für sämtliche Kausalitätstheorien. Auf welche Weise genau die Platitüden in einer Theorie umgesetzt werden ist individuell unterschiedlich. Außerdem gilt es laut Psillos dabei folgendes zu berücksichtigen:

(...) there are two central intuitions about causation that also need to be taken into account:

- The *intrinsic-relation* intuition: whether or not a sequence of two distinct events *c* and *e* is causal depends wholly on the events *c* and *e* and their own properties and relations, that is it depends wholly on

³ Psillos, Stathis: *Causation & Explanation*, Acumen, 2002, S. 6.

the intrinsic and local features of the actual sequence of events. For instance, according to this intuition, when we say that *the hitting with the hammer caused the smashing of the porcelain vase* what makes our assertion true has only to do with the properties of the particular hammer, the particular vase and the particular hitting.

- The *regularity* intuition: whether or not a sequence of two distinct events *c* and *e* is causal depends on whether or not events like *c* are regularly followed by events like *e*. This intuition is captured by the dictum “same cause, same effect” and is underpinned by an epistemic consideration; namely, that we are unwilling to pronounce a sequence of events *c* and *e* causal, unless there has been a regular association between events like *c* and events like *e*. For instance, according to this intuition, when we say that *the hitting with the hammer caused the smashing of the porcelain vase* what makes our assertion true has to do with the fact that the hitting of porcelain vases with hammers is regularly followed by the smashing of the vases.⁴

Die beiden Intuitionen hinsichtlich Kausalität nennt Psillos zum einen die intrinsische-Relation-Intuition und zum anderen die Regularitäts-Intuition. Die intrinsische-Relation-Intuition besagt im Wesentlichen, dass Verursachungsbeziehungen grundsätzlich intrinsischer Natur sind – d.h. ausschließlich die einzelnen Relata einer Beziehung sind für die Frage relevant, ob es sich dabei um eine Kausalbeziehung handelt oder nicht. Um zu beurteilen, ob eine Abfolge von Ereignissen kausal ist, genügt es also sich in der Analyse dieser Fragestellung auf die lokalen Eigenschaften der involvierten Ereignisse zu konzentrieren. So sind die Wahrmacher des Satzes ‚Das Schlagen mit dem Hammer hat das Zerschlagen der Porzellanvase verursacht‘ zum Beispiel ausschließlich in den Eigenschaften exakt dieses Schlagens, eben jener Vase und genau jenes Hammers zu suchen. Gemäß der Regularitäts-Intuition hingegen entscheidet letztendlich ein Vergleich mit anderen Fällen darüber, ob eine Abfolge von Ereignissen als kausal einzustufen ist oder nicht. Bei der Untersuchung zweier Ereignisse *c* und *e* wird geprüft, ob Ereignisse des Typs *C* regelmäßig Ereignisse des Typs *E* nach sich ziehen. Der zentrale Gedanke, der sich hinter dieser Intuition verbirgt, lässt sich in der Devise ‚Gleiche Ursache, gleiche Wirkung‘ zusammenfassen. Dieses Prinzip wiederum speist sich aus der allgemeinen Skepsis der Sequenz zweier Ereignisse *c* und *e* den Status einer Kausalbeziehung zuzuschreiben, solange nicht *C*-Ereignisse fortwährend *E*-Ereignisse zur Folge haben. Auf das obige Beispiel bezogen, bedeutet dies, dass es sich genau dann um eine kausale Abfolge handelt, wenn Hammerschläge für gewöhnlich Porzellanvasenbrüche bewirken.

Es ist offensichtlich, dass zwischen der intrinsische-Relation-Intuition und der Regularitäts-Intuition ein Spannungsverhältnis besteht. Dazu Psillos:

Now these two intuitions pull in contrary directions. The regularity intuition implies that a sequence of events is causal if and only if it instantiates a regularity. Hence, it implies that the relation of cause and effect is *extrinsic* to its relata. It makes causation dependent on *general* facts; on what happens at other places and at other times. The intrinsic-relation intuition opposes all this. It takes causation to be wholly dependent on *singular* facts: on what happens there and then, in the actual sequence of events, independently of any regularities. It would be a daunting (not to say outright impossible) task to advance a theory that respects both of these intuitions. Most

⁴ Psillos, Stathis: *Causation & Explanation*, Acumen, 2002, S. 6f.

typically, Humeans base their theories on the regularity intuition, while non-Humeans base theirs on the intrinsic-relation one.⁵

Laut Psillos wäre es ein schwieriges – um nicht zu sagen geradezu unmögliches – Unterfangen, beide Intuitionen in einer Theorie zu vereinen. Denn offensichtlich explizieren sie entgegengesetzte und vermeintlich unvereinbare Aspekte der vortheoretischen Vorstellung von Kausalität. Im Sinne der Regularitäts-Intuition sind Kausalbeziehungen nämlich ihrem Wesen nach extrinsisch. Schließlich hängen sie von allgemeinen Tatsachen ab, insbesondere von der Frage, ob C-Ereignisse stets E-Ereignisse auslösen. Für die Bestimmung, ob zwei Ereignisse kausal verbunden sind, spielt es also eine Rolle was anderswo und zu anderen Zeitpunkten geschieht. Kurzum, die Kausalrelation ist sonach extrinsisch im Hinblick auf ihre Relata. Gewissermaßen die Antithese hierzu bildet die intrinsische-Relation-Intuition, der zufolge Verursachung eine rein intrinsische Relation ist. Um festzustellen ob die Abfolge von zwei Ereignissen kausal verknüpft ist, genügt es folglich nur diese beiden Ereignisse und deren entsprechende Eigenschaften zu betrachten. In anderen Worten sind ausschließlich singuläre Fakten für diese Beurteilung von Bedeutung – unter der Leitfrage: was passiert hier und jetzt bei genau dieser Chronologie von Ereignissen? Eine gewisse Einordnung der Diskussion über Kausalität lässt sich bereits anhand dieser beiden Intuitionen vornehmen: während Humeaner sich grundsätzlich an der Regularitäts-Intuition orientieren, bauen Nicht-Humeaner ihre Theorien meist auf der intrinsischen-Relation-Intuition auf.

Für eine feinkörnigere Einteilung der Debatte, unterteilt Psillos in:

(...) three dimensions along which the discussion about causation can be based. (...) The first concerns the distinction between generalist and singularist theories. The second dimension concerns the distinction between theories that aim at an extrinsic characterization of causal relations and theories that go for an intrinsic one. The third dimension concerns the distinction between reductive approaches and non-reductive ones. Reductive approaches argue that causation is dependent on (some say it supervenes on) non-causal features of the world (e.g. regularities), while non-reductive accounts take causation to be ontically autonomous: an irreducible relation among events. On a first approximation, then, one could say that Humean accounts of causation take the first sides of the three dimensions: they are generalist, extrinsic and reductive. And non-Humean accounts take at least one of the remaining sides of the three distinctions: they are singularist or intrinsic or non-reductive.⁶

Die drei Kategorien, in die sich Kausalitätstheorien untergliedern lassen, können Psillos zufolge anhand der folgenden drei Gegensatzpaare veranschaulicht werden: generalistisch – singulär, extrinsisch – intrinsisch und reduktiv – nicht-reduktiv. Während generalistische Ansätze von Kausalität davon ausgehen, dass Verursachung von allgemeinen Mustern, wie z.B. von Regularitäten, abhängig ist, fokussieren singuläre Modelle ausschließlich auf Abfolgen von Einzelereignissen. Zwar ist es auch nach einem generalistischen Kausalitätsverständnis

⁵ Psillos, Stathis: *Causation & Explanation*, Acumen, 2002, S. 7.

⁶ Psillos, Stathis: *Causation & Explanation*, Acumen, 2002, S. 7f.

möglich eine individuelle Sequenz von Ereignissen als kausal zu klassifizieren, dies ist allerdings allein der Zugehörigkeit der entsprechenden Ereignisse zu einer bestimmten Menge von Ereignissen geschuldet. Im Gegensatz dazu argumentieren Vertreter von singulären Kausalitätstheorien, dass es für die Frage, ob eine Abfolge von zwei Ereignissen *c* und *e* kausal ist, keine Rolle spielt wie sich ähnliche Ereignisse in der Vergangenheit verhalten haben oder wie sie sich in der Zukunft verhalten werden.⁷ Die zweite Dimension, die Psillos bei der Strukturierung der Kausalitätsdebatte heranzieht, unterscheidet Theorien, die Kausalität als eine intrinsische Relation verstehen und solchen, die sie als extrinsische Beziehung auffassen. Dabei gilt eine Beziehung genau dann als intrinsisch, wenn sie lediglich auf dem Verhältnis der jeweiligen Relata zueinander basiert. Infolgedessen ist eine Relation als extrinsisch einzustufen, wenn darüber hinaus noch andere Faktoren von Bedeutung sind. Zur Veranschaulichung dieser Begriffsdifferenzierung führt Psillos folgende Beispiele an: *x hat mehr Masse als y* ist demnach eine intrinsische Relation. Denn setzt man beispielsweise die Paarung <Sonne; Erde> ein, so stellt sich heraus, dass der Wahrheitsgehalt der Aussage ‚Die Sonne hat mehr Masse als die Erde‘ allein davon abhängt wie Erde und Sonne wechselseitig in Beziehung stehen. Hierbei ist es völlig unerheblich wie sich Erde und Sonne zu irgendetwas anderem im Universum verhalten. Bei extrinsischen Relationen hingegen – wie z.B. bei *x und y gehören demselben Besitzer* – hängt die Wahr- bzw. Falschheit einer Aussage nicht ausschließlich von der Beziehung von *x* und *y* ab. Vielmehr ist hier die Verbindung von *x* und *y* zu etwas Drittem (nämlich ihrem jeweiligen Besitzer) von entscheidender Bedeutung.⁸ Offensichtlich besteht ein enger Zusammenhang zwischen Psillos’ erstem und zweitem Unterscheidungsmerkmal von Kausalitätstheorien. Schließlich verstehen singuläre Ansätze Kausalität in der Regel als intrinsische Relation, wohingegen generalistische Modelle Verursachung extrinsisch interpretieren. Das hängt damit zusammen, dass gemäß des generalistischen Verständnisses *c* *e* verursacht, weil eine Regularität zwischen C- und E-Ereignissen besteht. Somit ist nicht exklusiv die Beziehung zwischen *c* und *e* ausschlaggebend, sondern auch deren Verhältnis zu etwas Anderem – nämlich der Regularität. Aus diesem Grund ist das Kausalitätsverständnis generalistischer Ansätze zwangsläufig extrinsisch. Umgekehrt verpflichtet eine singuläre Auffassung von Verursachung nicht unbedingt auf ein intrinsisches Verständnis, so Psillos. So kann beispielsweise die Handlungstheorie von Menzies und Price durchaus als singuläre Kausalitätstheorie verstanden werden; nichtsdestotrotz handelt es sich hier allerdings um eine extrinsische Auffassung von Verursachung. Die zentrale These dieser Handlungstheorie besagt, dass ein Ereignis *c* genau dann die Ursache eines anderen Ereignisses *e*

⁷ Vgl. Psillos, Stathis: *Causation & Explanation*, Acumen, 2002, S. 127.

⁸ Vgl. Psillos, Stathis: *Causation & Explanation*, Acumen, 2002, S. 128.

ist, wenn das Hervorbringen des Auftretens von c ein geeignetes Mittel für einen frei Handelnden wäre, um ein Stattfinden von e herbeizuführen.⁹ Demnach ist es also durchaus möglich diese Auffassung als singuläre Vorstellung von Kausalität zu interpretieren, denn im Hinblick auf die Frage, ob das Verursachen von c für einen Handelnden ein Mittel zum Bewirken von e darstellt, spielt es keine Rolle, ob eine Regularität zwischen C-Ereignissen und E-Ereignissen besteht. Allerdings macht diese Bezugnahme auf etwas Drittes – nämlich einen frei Handelnden – dieses Kausalitätsverständnis auch zu einem extrinsischen.¹⁰ Das dritte und letzte Unterscheidungsmerkmal von Kausalitätstheorien bezieht sich auf den ontologischen Status von Verursachungsbeziehungen. Die Kernfrage lautet: Ist Kausalität ontologisch autonom oder hängt sie von nicht-kausalen Faktoren ab? Grundsätzlich wird ontologische Abhängigkeit, so Psillos, auf zwei unterschiedliche Weisen artikuliert: als vollständige Reduktion oder als Supervenienz. Bei der kompletten Reduktion werden die Wahrheitsbedingungen von Kausalaussagen zur Gänze von nicht-kausalen Merkmalen bestimmt. Im Gegensatz dazu handelt es sich bei Supervenienz um eine schwächere Form der ontologischen Abhängigkeit. Zwar gibt es mehrere Spielarten der Supervenienz, prinzipiell ist ihnen allen aber folgender Grundgedanke gemein: wenn zwei (mögliche) Welten im Hinblick auf ihre nicht-kausalen Fakten identisch sind, dann sind diese beiden Welten auch bezüglich ihrer kausalen Fakten identisch. D.h. die nicht-kausalen Tatsachen legen die kausalen Tatsachen in vollem Umfang fest. Sind sämtliche kausale Gegebenheiten auf nicht-kausale Sachverhalte reduzierbar, impliziert dies im Grunde eine Identität zwischen beiden Faktenmengen, supervenieren diese jedoch auf jenen, dann sind die beiden Mengen hingegen nicht zwangsläufig identisch. Nichtsdestotrotz behaupten Anhänger von Supervenienztheorien, dass die subvenienten Begebenheiten die supervenienten determinieren. Eine Schlüsselrolle spielt dabei die raum-zeitliche Verteilung der lokalen Eigenschaften, ist diese nämlich festgelegt, dann ist damit auch alles andere bestimmt. In Anbetracht der anderen Kategorien lässt sich festhalten, dass sich unter den supervenienzbasierten Theorien sowohl singuläre als auch generalistische, genauso wie intrinsische und extrinsische Ansätze finden.

Kausalitätstheorien

Um sich einen Überblick über die Hauptströmungen in der zeitgenössischen Kausalitätsdebatte zu verschaffen, bietet es sich an zunächst einen radikalen Kritiker vom Konzept der Verur-

⁹ Vgl. Psillos, Stathis: *Causation & Explanation*, Acumen, 2002, S. 102.

¹⁰ Vgl. Psillos, Stathis: *Causation & Explanation*, Acumen, 2002, S. 129.

sachung zu Wort kommen zu lassen. Zu Beginn des 20. Jahrhunderts fällt Bertrand Russell in seinem Aufsatz *On the Notion of Cause* folgendes vernichtende Urteil über Ursachen und Kausalität: „(...) the reason why physics has ceased to look for causes is that, in fact, there are no such things. The law of causality, I believe, like much that passes muster among philosophers, is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm.“¹¹ Russell vertritt also den Standpunkt, dass es so etwas wie Ursachen im Grunde gar nicht gibt. Dementsprechend ist auch die Suche nach ihnen ein sinnloses Unterfangen. Die grundlegendste aller Wissenschaften, die Physik, hat dies bereits erkannt und nun ist es an der Zeit, dass es die Philosophie ihr gleichtut, so Russell. Schließlich ist Kausalität nichts anderes als ein Rudiment längst vergangener Tage und der Begriff ‚Ursache‘ sollte gänzlich aus dem (philosophischen) Vokabular getilgt werden. In seiner Rezeption Russells, listet Max Kistler drei Argumente auf, die Russell zur Untermauerung seiner These heranzieht:

Russell's first argument is that the idea, inherent in the concept of causation, of an intrinsic difference between an active cause and a passive effect, is anthropomorphic. This makes the concept inadequate for the analysis of natural causal relations, with the exception of human actions.(...)

His second argument will prove particularly important in what follows. The concept of causation is associated with the "principle of causality" according to which causality is the source of regularities in nature. (...) The principle of causality says that the same cause always has the same effect. Now, it is essential to the concept of causation to be applicable to macroscopic events such as car accidents, avalanches, or deaths. The formidable complexity of such events contrasts with the simplicity of the conditions that are the object of the laws of physics. According to Russell, the requirements of applicability to complex macroscopic events and of following the strict regularity of a law of nature, are incompatible, although both are central to causation. The more the description of a complex event involves sufficient detail and completeness to allow the application of physical laws, the less it is possible that the event will ever recur. However, without recurrence, the idea of regularity is unintelligible. This blocks the possibility to justify the scientific legitimacy of the concept of causation by its link to lawful regularity via the principle of causality, which might have appeared promising, given that modern science has substituted the search for laws for the search for causes.

Russell's second argument establishes that there are no regularities at the level of concrete events. His third argument goes a step further. It says that in 20th century physics, there are not any laws of succession at all, which might have played a role of grounding causation. The progress of physics consists in replacing them by laws of coexistence or by equivalences or functional dependencies expressed in equations. This thesis about the form of laws in mature science explains according to Russell the failure of the project to make the concept of causation respectable by reducing it to nomological regularity.¹²

Russells erster Kritikpunkt am Konzept der Kausalität setzt an der Unterscheidung zwischen einer aktiven Ursache einerseits und einer passiven Wirkung andererseits an. Diese Vorstellung ist laut Russell zu stark an den Menschen angelehnt und eigne sich daher ausschließlich zur Analyse menschlichen Verhaltens und nicht zur Untersuchung anderer Kausalbeziehungen in der Natur. Der zweite Einwand Russells zielt auf einen integralen Bestandteil der Idee von Kausalität – nämlich das Gesetz der Kausalität, demzufolge Verursachung für Regularität-

¹¹ Russell, Bertrand: „On the Notion of Cause“ in: Russell, Bertrand: *Mysticism and Logic and other Essays*, 1951, S. 180.

¹² Kistler, Max: „Causation in contemporary analytical philosophy“, in: *Quaestio*, Vol.2, 2002, S. 635f.

ten in der Natur verantwortlich ist. Gemäß dieses Prinzips hat dieselbe Ursache stets dieselbe Wirkung. Allerdings scheint es, als wären die Zuordnung zu Regelmäßigkeiten und das Fallen unter strikte Naturgesetze unvereinbar. Denn um Ereignisse in gewisse Regularitäten einzuordnen ist es notwendig sich auf makroskopische Ereignisse, wie z.B. Autounfälle, Lawinen oder Todesfälle, zu beziehen. Die Vielschichtigkeit solcher Ereignisse widerspricht jedoch der Einfachheit der Bedingungen, die beim Ausformulieren von Naturgesetzen eingefordert wird. Schließlich steigt mit zunehmender Komplexität der Beschreibung eines Ereignisses die Wahrscheinlichkeit, dass sich dieses Vorkommnis niemals exakt so wiederholt. Kurzum, auf der Ebene von konkreten Ereignissen gibt es keine Regelmäßigkeiten. Russell entwickelt sein drittes Argument ausgehend von der Form von Naturgesetzen. Diese hat sich nämlich in der Physik des 20. Jahrhunderts maßgeblich geändert: während man vorher von Abfolge-Gesetzen ausgegangen ist, ist man inzwischen dazu übergegangen Koexistenz-Gesetze anzunehmen, die mittels Äquivalenzen oder funktionaler Abhängigkeiten dargestellt werden. Letztere erlauben jedoch keine Reduktion von Kausalität auf nomologische Regularität. Schließlich ist es mit Äquivalenzen und funktionalen Abhängigkeiten unmöglich zwei Kerngedanken von Kausalität darzustellen. Diese sind zum Einen, dass die Ursache *zeitlich vor* der Wirkung liegt und zum Anderen, dass Kausalität eine *asymmetrische* Beziehung ist, weil die Ursache die Wirkung herbeiführt und nicht umgekehrt.

Allerdings blieb Russells Forderung nach einem Bann jeglichen kausalen Vokabulars aus dem Wortschatz der Philosophie ohne Erfolg, stattdessen fand die „(...) reintegration of the notion of causation within the conceptual apparatus of the philosophy of science (...) its canonical expression in the assimilation of causation to the relation of explanation, conceived according to the so-called ‘deductive-nomological’ model.“¹³ Dem deduktiv-nomologischen Modell, oder kurz D-N-Modell, zufolge ist eine wissenschaftliche Erklärung im Grunde ein deduktives Argument. Die Konklusion des Arguments entspricht dem Explanandum – also dem Phänomen, das es zu erklären gilt. Die Prämissen in ihrer Gesamtheit repräsentieren das Explanans – somit dasjenige, das die Erklärungsleistung trägt. Das Explanans setzt sich aus den Anfangsbedingungen und den allgemeinen Gesetzen, die auf den entsprechenden Fall zutreffen, zusammen. Zudem gelten für eine erfolgreiche Erklärung gewisse Adäquatheitsbedingungen: erstens muss das Explanandum mit logischer Notwendigkeit aus dem Explanans folgen. Zweitens muss das Explanans einen wesentlichen Bezug zu allgemeinen Gesetzen aufweisen. Drittens leistet das D-N-Modell ausschließlich Erklärungen empirischer Phänomene und deshalb muss das Explanans einen empirischen – d.h. falsifizierbaren – Gehalt besitzen.

¹³ Kistler, Max: „Causation in contemporary analytical philosophy“, in: *Quaestio*, Vol.2, 2002, S. 638.

Viertens müssen alle Aussagen im Explanans wahr sein, damit auch die Erklärung insgesamt wahr sein kann.¹⁴

Im Hinblick auf die Frage nach dem Verhältnis von Erklärung und Kausalität gelangen die beiden Väter des D-N-Modells wissenschaftlicher Erklärung laut Kistler zu folgendem Schluss:

According to Hempel and Oppenheim, any explanation that satisfies the requirements of the D-N-model is *ipso facto* a causal explanation. Hence, it is possible to use the model to analyse the causal relation: the conjunction of the initial conditions is taken to be a description of the cause and the event described by the explanandum is taken to be its effect. Insofar as the laws appearing in the explanans are explicitly considered to be empirical regularities, the assimilation of the cause-effect relation to the explanatory relation in the D-N-model leads to a version of the regularity theory of causation.¹⁵

Nach Hempel und Oppenheim ist eine Erklärung anhand des D-N-Modells per se eine kausale Erklärung. Daher eignet sich dieses Erklärungsmodell auch zur Untersuchung von Kausalbeziehungen. In diesem Zusammenhang kommt die Summe aller Anfangsbedingungen einer Beschreibung der Ursache gleich und das im Explanandum beschriebene Ereignis verkörpert die Wirkung. Da die allgemeinen Gesetze, auf die im Explanans Bezug genommen wird, ausdrücklich empirische Regelmäßigkeiten widerspiegeln, fällt die Analyse von Kausalbeziehungen im Rahmen des D-N-Modells unter die Regularitätstheorien von Kausalität.

Im weiteren Verlauf seines Aufsatzes analysiert Kistler vier unterschiedliche Typen von Kausalitätstheorien unter diesen Gesichtspunkten. Diese sind kontrafaktische, interventionistische, probabilistische und schließlich Transfer-Theorien.¹⁶

David Lewis ist einer der Hauptverfechter einer *kontrafaktischen* Auffassung von Kausalität in der zeitgenössischen Philosophie. Die Kernidee seiner Theorie besteht in:

(...) reducing causation between two events to a finite chain of intermediate events, each link of which causally depends on the preceding. The notion of *causal dependence* between events is then analysed with the help of the *counterfactual* dependence between the propositions saying that those events have taken place. According to the counterfactual theory of causation, event *e* depends causally on event *c* if the following two counterfactuals are true: if *c* had occurred, *e* would have occurred; if *c* had not occurred, *e* would not have occurred. The distinction between causal dependence and the causal relation itself is introduced in order to guarantee the transitivity of the causal relation, for causal dependence is not transitive.¹⁷

Zunächst gilt es laut Lewis also die Verursachungsbeziehung zwischen zwei Ereignissen als eine endliche Kette intermediärer Ereignisse aufzufassen, von denen jedes Glied vom vorausgehenden kausal abhängig ist. Der Begriff der kausalen Dependenz zwischen Ereignissen

¹⁴ Vgl. Kistler, Max: „Causation in contemporary analytical philosophy“, in: *Quaestio*, Vol.2, 2002, S. 639.

¹⁵ Kistler, Max: „Causation in contemporary analytical philosophy“, in: *Quaestio*, Vol.2, 2002, S. 640.

¹⁶ Vgl. Kistler, Max: „Causation in contemporary analytical philosophy“, in: *Quaestio*, Vol.2, 2002, S. 645.

¹⁷ Kistler, Max: „Causation in contemporary analytical philosophy“, in: *Quaestio*, Vol.2, 2002, S. 645.

wird anschließend mithilfe des Konzepts der kontrafaktischen Abhängigkeit zwischen den Aussagesätzen, die behaupten diese Ereignisse haben stattgefunden, beleuchtet. Dabei ist ein Ereignis *e* kontrafaktisch von einem Ereignis *c* abhängig, wenn folgende kontrafaktische Konditionale wahr sind. Wenn *c* eingetreten wäre, wäre auch *e* eingetreten. Wenn sich *c* nicht ereignet hätte, wäre auch *e* ausgeblieben. Lewis' Motivation für die Differenzierung von kausaler Abhängigkeit auf der einen Seite und der Kausalrelation selbst auf der anderen Seite, gilt der Absicherung von Transitivität. Denn während Transitivität¹⁸ zumeist als ein wesentliches Charakteristikum von Verursachungsbeziehungen dargestellt wird, sind sowohl die Beziehung kausaler als auch kontrafaktischer Abhängigkeit nicht zwangsläufig transitiv. So ist es beispielsweise durchaus möglich, dass folgendes Argument von wahren Prämissen zu einer falschen Konklusion führt:

- P1 Wenn Otto zu der Feier gegangen wäre, dann wäre Anna auch gegangen.
- P2 Wenn Anna zu der Feier gegangen wäre, dann wäre Waldo auch gegangen.
- K Wenn Otto zu der Feier gegangen wäre, dann wäre Waldo auch gegangen.

Dies wäre z.B. dann ein Fehlschluss, wenn Otto und Anna ein Paar sind, wohingegen Waldo unglücklich in Anna verliebt ist und daher Begegnungen mit seinem Rivalen Otto um jeden Preis zu vermeiden versucht.¹⁹

Eines der Hauptprobleme sowohl der D-N-Theorie von Kausalität als auch der kontrafaktischen Auffassung von Verursachung ist eine angemessene Abbildung der Asymmetrie von Kausalbeziehungen. Die Strategie von *interventionistischen* Kausalitätstheorien für den Umgang mit diesen Schwierigkeiten ist „(...) to overcome them by grounding that asymmetry in the asymmetry between the past, which we cannot influence, and the future, which presents itself to us as agents as an «open 'space' of possibilities».“²⁰ Die Asymmetrie zwischen Ursache und Wirkung ist also in erster Linie der Gerichtetheit der Zeit geschuldet. Ebenso wenig wie die Vergangenheit von der Zukunft abhängen kann, ist es auch nicht möglich, dass eine Ursache von ihrer eigenen Wirkung bedingt wird. Ein weiteres Wesensmerkmal interventionistischer Kausalitätstheorien ist die starke Bezugnahme auf menschliche Akteure. So kann ein Handelnder die Vergangenheit nicht verändern, wohingegen sich ihm die Zukunft als ein freier Möglichkeitsraum offenbart. Entsprechend kann die Wirkung einer Handlung nicht ihre Ursache herbeiführen, denn jene tritt stets zeitlich vor dieser ein. Die zentrale Botschaft der

¹⁸ Wenn *a* *b* verursacht und *b* wiederum *c* verursacht, dann ist – falls die Kausalrelation transitiv ist – damit *a* auch notwendigerweise die Ursache von *c*.

¹⁹ Vgl. Kistler, Max: „*Causation in contemporary analytical philosophy*“, in: *Quaestio*, Vol.2, 2002, S. 645f.

²⁰ Kistler, Max: „*Causation in contemporary analytical philosophy*“, in: *Quaestio*, Vol.2, 2002, S. 653.

interventionistischen Kausalitätsauffassung ist laut Kistler diese: „More generally, the interventionist or manipulability account of causation reduces the existence of a causal relation between two distinct events A and B to the fact that «bringing about the occurrence of A would be an effective means by which a free agent could bring about the occurrence of B».“²¹ Es geht also im Wesentlichen darum die Möglichkeit einer Kausalbeziehung zwischen zwei Ereignissen A und B daran festzumachen, dass das Hervorbringen von A für einen frei Handelnden ein gutes Mittel wäre, um das Eintreten von B hervorzurufen.

Probabilistische Kausalitätstheorien vereint laut Kistler folgender Grundgedanke: „The fundamental thesis of the probabilistic approach to causation is that A is cause of B if and only if the occurrence of A raises the chance of the occurrence of B.“²² Wenn also das Eintreten eines Ereignisses A die Wahrscheinlichkeit des Stattfindens eines Ereignisses B erhöht, kann nach probabilistischem Kausalitätsverständnis A als Ursache von B aufgefasst werden. Dies offenbart aber unmittelbar eine wesentliche Schwachstelle dieses Ansatzes: „The major problem with this approach concerns its application to the analysis of singular causation, i.e. to causal relations between particular pairs of events.“²³ Die Hauptschwierigkeit der probabilistischen Kausalitätsauffassung ist nach Kistler also ihre Anwendung auf Einzelfälle. Nimmt man beispielsweise die Annahme, dass Rauchen die Wahrscheinlichkeit einer Erkrankung an Lungenkrebs erhöht. So scheint eine statistische Korrelation zwischen Nikotinkonsum und Krebsleiden allein noch kein kausales Urteil für den Einzelfall zu erlauben. Denn es ist etwa durchaus möglich, dass ein Raucher nicht an Lungenkrebs erkrankt, oder, dass ein Tumor in der Lunge eines Rauchers wächst, weil der Betroffene zugleich Asbest ausgesetzt war. Offensichtlich rechtfertigt ein statistischer Zusammenhang noch nicht die Zuschreibung einer kausalen Beziehung, denn:

The statistical correlations that constitute the point of departure of the probabilistic analysis are often spurious, in the sense that the correlation exists between two (types of) effects of a common (type of) cause, but where these effects are not linked as cause and effect. Hence, the theory needs to be completed by the following requirement: the fact that A raises the probability of B justifies the belief in the existence of a causal relation between (events of type) A and B, only if there is no third factor that is the real cause of both. Such a factor is called a ‘screening factor’.²⁴

Den Ausgangspunkt für probabilistische Kausalitätstheorien bilden statistische Korrelationen, diese liefern jedoch gelegentlich ein Zerrbild der Situation. Denn es ist beispielsweise durchaus möglich, dass zwei Wirkungen einer gemeinsamen Ursache miteinander korrelieren, aber

²¹ Kistler, Max: „Causation in contemporary analytical philosophy“, in: *Quaestio*, Vol.2, 2002, S. 653.

²² Kistler, Max: „Causation in contemporary analytical philosophy“, in: *Quaestio*, Vol.2, 2002, S. 655.

²³ Kistler, Max: „Causation in contemporary analytical philosophy“, in: *Quaestio*, Vol.2, 2002, S. 655.

²⁴ Kistler, Max: „Causation in contemporary analytical philosophy“, in: *Quaestio*, Vol.2, 2002, S. 657.

nicht in einem Ursache-Wirkungs-Verhältnis zueinander stehen. Daher müssen probabilistische Verursachungstheorien eine Einschränkung – nämlich eine Abschirmregel (screening factor) – einführen, so dass gilt: die Tatsache, dass A die Wahrscheinlichkeit von B erhöht, gestattet die Vermutung, dass eine kausale Beziehung zwischen den Ereignistypen A und B besteht – zumindest solange beide nicht von einem dritten Ereignistyp (d.h. einer gemeinsamen Ursache) ausgelöst werden .

Wie problematisch die Zuschreibung eines Kausalverhältnisses aufgrund einer statistischen Wahrscheinlichkeitsverteilung sein kann, verdeutlicht das so genannte Simpson-Paradox. Das Simpson-Paradox zeigt auf, dass bei der Analyse der Wahrscheinlichkeitsverteilung entscheidend ist in welche Untergruppen man eine zu untersuchende Menge einteilt. Denn je nach Aufteilung ist eine unterschiedliche, teilweise sogar widersprüchliche, Interpretation der Datenlage möglich. Als Beispiel soll die vermeintliche Diskriminierung von Frauen bei der Bewerberauswahl der Universität Berkeley dienen. In folgender Aufstellung ist das jeweilige Verhältnis von Bewerber zu aufgenommenem/r Student/in je nach Disziplin dargestellt:

	Fachbereich 1	Fachbereich 2	Gesamt
Männlich	81/90	2/10	83/100
Weiblich	9/10	18/90	27/100

Die absoluten Werte der Zulassungszahlen legen eine deutliche Benachteiligung von Frauen in der Aufnahmepolitik der Universität nahe, so erhalten 83% der männlichen Bewerber einen Studienplatz, wohingegen nur 27% der weiblichen Antragstellerinnen einen positiven Studienplatzbescheid erhalten. Betrachtet man ausschließlich diese Zahlen liegt ein kausaler Zusammenhang zwischen den Eigenschaften männlich zu sein und bei der Aufnahme begünstigt zu werden nahe. Gliedert man die Gruppe der Studienplatzbewerber jedoch noch weiter nach den Fachbereichen auf, in die sie aufgenommen werden möchten, so ergibt sich plötzlich ein völlig anderes Bild: 90% der Interessenten erhalten im Fachbereich 1 das Angebot auf einen Studienplatz, im Fachbereich 2 liegt die Aufnahmequote bei 20% - und zwar unabhängig vom Geschlecht. Mit der Kenntnis dieser Prozentwerte muss die ursprüngliche Vermutung, die eine Bevorteilung der männlichen Studienanwärter behauptet hat, wieder verworfen werden. Vielmehr scheint an anderer Stelle ein kausaler Zusammenhang zu bestehen: männliche Bewerber haben offenbar eine Neigung sich für Studienplätze im leichter zugänglichen Fachbe-

reich 1 zu bewerben, wohingegen weibliche Kandidatinnen eher in den Fachbereich 2, der im Allgemeinen deutlich schwerer zugänglich ist, aufgenommen werden möchten.²⁵

Aufgrund dieser Schwierigkeiten zieht Kistler folgende Bilanz:

In short, it is impossible to draw any conclusive causal information whatsoever from an analysis of conditional probabilities calculated in an inhomogeneous set. To the extent that it is possible to introduce new partitions by paying attention to new factors, it can always happen that the conditional properties change radically. The fact that judgements of statistical correlation always remain relative to the choice of partition, and that the direction of correlation may depend on that choice, makes them inappropriate to ground absolute causal judgements.²⁶

Die Analyse von bedingten Wahrscheinlichkeiten in einer heterogenen Menge erlaubt nach Kistler also keinerlei Rückschlüsse auf kausale Beziehungen zwischen ihren Elementen. Schließlich sei es möglich, dass sich die bedingten Wahrscheinlichkeiten durch eine andere Gruppierung oder durch eine Fokussierung auf neue Aspekte fundamental verändern. Die Bewertung der statistischen Korrelation hängt folglich im Wesentlichen davon ab in welche Segmente die zu analysierende Datenmenge unterteilt wird. Die Tatsache wiederum, dass diese Auswahl darüber bestimmt, ob die Elemente positiv oder negativ miteinander korrelieren, scheint die Urteile statistischer Korrelationen als Grundlage kausaler Urteile zu disqualifizieren.

Die Methode der Wahl für Verfechter einer *Transfer-Theorie* von Kausalität ist folgende Vorgehensweise:

(...) to reduce the causal relation between two particular events to the transmission between them of some amount of a CQ [conserved quantity, S.H.]. It is up to science to establish the explicit list of all the CQ that exist in nature and that can therefore ground causal relations.

This idea can be elaborated so as to provide a satisfactory account of the situations that raise insuperable problems for the theories we examined earlier. The first qualification concerns the asymmetry of causation. The concept of transference is symmetrical: the individual amount of CQ X is transferred from *a* to *b* if it is present both in event *a* and in event *b*. In a possible world in which all processes are reversible, there is transmission and therefore causation, but causal relations (and with them, time, if the direction of time is reduced to the direction of causation) are symmetric. In such a world there is no intrinsic difference between two terms of a causal relation, which makes one the cause rather than the other. However, in the actual world, most processes are irreversible. This irreversibility is at the origin of the asymmetry of both causation and time. Borrowing from Reichenbach the concept of a causal net in which all events are causally linked to other events, so that no event is causally absolutely isolated, it is possible to consider that insofar as a large majority of the processes in the net possess the same intrinsic direction, which is the direction of growing entropy, that direction is the direction of the entire net. In this way, it is possible to attribute a direction even to intrinsically symmetrical processes, i.e. reversible processes.

The problem of effects can then be solved by appealing to the empirical fact that the majority of actual causal processes have the same intrinsic direction: *a* is cause of *b* because first, an individual amount of a CQ is transmitted between them and second, because either the transmission process between *a* and *b* is physically asymmetrical and irreversible or it is reversible but part of a network of causal processes possessing a global direction and imposing it on the transmission process between *a* and *b*.

²⁵ Gesamter Absatz, vgl. Kistler, Max: „Causation in contemporary analytical philosophy“, in: *Quaestio*, Vol.2, 2002, S. 657f.

²⁶ Kistler, Max: „Causation in contemporary analytical philosophy“, in: *Quaestio*, Vol.2, 2002, S. 657.

The transference theory overcomes the problems of epiphenomena, preemption and causal overdetermination rather simply.²⁷

Zunächst geht es bei Transfer-Theorien von Kausalität also darum aufzuzeigen, dass sich die Verursachungsrelation zwischen zwei Ereignissen auf die Übertragung einer Erhaltungsgröße zurückzuführen lässt. Wobei die Identifizierung des Katalogs aller Erhaltungsgrößen in den Aufgabenbereich der empirischen Wissenschaften fällt. Für die oben beschriebenen Schwierigkeiten²⁸ von Verursachungstheorien haben Transfer-Theorien innovative Lösungsansätze. Auf das Problem der Asymmetrie von Kausalbeziehungen finden Transfer-Ansätze eine interessante Antwort, obwohl die Übertragungsrelation selbst symmetrisch ist. Schließlich ist die Erhaltungsgröße per definitionem sowohl im Ursachen- als auch im Wirkungsereignis in gleichem Maße vorhanden. Die Wurzel der Asymmetrie der Verursachungsbeziehung ist nach Transfer-Theorien in der Struktur der aktuellen Welt zu suchen. Genauer gesagt, in der prinzipiellen Unumkehrbarkeit der meisten Prozesse. Allerdings sind durchaus mögliche Welten denkbar, in denen alle Prozesse grundsätzlich umkehrbar sind. In solchen möglichen Welten verflüchtigt sich auch jeglicher intrinsische Unterschied zwischen Ursache und Wirkung und Kausalität wäre unter diesen Bedingungen eine symmetrische Beziehung. Angenommen alle Ereignisse sind in ein kausales Netz eingebettet, durch das sie miteinander verbunden sind. Eine Vielzahl der Ereignisse hat dieselbe intrinsische Richtung (i.e. hin zur größtmöglichen Entropie) und dadurch wird auch die Richtung des gesamten Netzes determiniert. Damit wäre es möglich, dass Prozesse intrinsisch symmetrisch sind und trotzdem eine Richtung besitzen. Das sogenannte ‚problem of effects‘ kann auf folgende Art gelöst werden: ein Ereignis a ist die Ursache eines Ereignisses b, weil erstens eine Erhaltungsgröße von a auf b übertragen wurde. Und zweitens ist die Übertragung zwischen a und b entweder physikalisch asymmetrisch und unumkehrbar. Oder sie ist umkehrbar und Teil eines Netzes kausaler Prozesse, das insgesamt in eine bestimmte Richtung steuert und diese auf den Transfer zwischen a und b überträgt. Die Probleme der Epiphänomene, der Präemption und der kausaler Überdeterminierung enträtseln Transfer-Theorien verhältnismäßig einfach.²⁹

²⁷ In Kistler, Max: „*Causation in contemporary analytical philosophy*“, in: *Quaestio*, Vol.2, 2002, S. 663f sind einige Beispiele für die genannten Schwierigkeiten zu finden.

²⁸ An dieser Stelle sind insbesondere die Asymmetrie von Kausalbeziehungen und das, was Lewis ‚problem of effects‘, ‚problem of epiphenomena‘ und ‚redundant causation‘ nennt, gemeint.

²⁹ Kistler, Max: „*Causation in contemporary analytical philosophy*“, in: *Quaestio*, Vol.2, 2002, S. 646f.

Zielsetzung der Arbeit und ausführliche Darstellung der Vorgehensweise

Ziel dieser Arbeit ist es die Grundzüge der umfangreichen und anhaltenden Diskussion zu den Schwierigkeiten, die einer adäquaten Vorstellung mentaler Verursachung entgegenstehen, historisch nachzuzeichnen und in ihre Hauptargumentationsstränge aufzugliedern. Dieser Problemaufriss soll dabei behilflich sein diese, zum Teil sehr unübersichtliche, Debatte innerhalb der Philosophie des Geistes klar zu strukturieren und zu vereinheitlichen. Dabei werde ich an erster Stelle die wesentlichen Elemente, die einer angemessenen Darstellung mentaler Kausalität zuwiderlaufen, beschreiben. Sodann werde ich auf Kritikpunkte an den einzelnen Argumenten eingehen, mögliche Lösungsstrategien zu diesen Einwänden aufzeigen und differenziert diskutieren. In diesem Zusammenhang werde ich auch einige Schlüsselpositionen detailliert vorstellen und kritisch beleuchten. Am Ende der Arbeit werde ich darlegen, dass die bislang beste Strategie im Umgang mit den skizzierten Problemen mentaler Kausalität Fred Dretskes Dual-Explanandum Theorie ist.

Grundsätzlich können zwei unterschiedliche Annahmen zu Problemen im Hinblick auf die Erklärung der kausalen Interaktion zwischen Mentalem und Physikalischem führen: einerseits die Hypothese von mentalen Substanzen, andererseits die Behauptung von eigenständigen mentalen Eigenschaften im Rahmen der Physik. Die Herausforderungen, für diejenigen, die mentale Substanzen in ihre Ontologie aufnehmen, werde ich unter der Überschrift *Interaktionsproblem* diskutieren. Die Probleme der Philosophen, die mentalen Eigenschaften eine gewisse Autonomie in einer im Grunde physischen Welt zuschreiben, habe ich in drei Kapitel unterteilt: das *Problem um die Anomalie des Mentalen*, das *Exklusionsproblem* und *Externalismus und Anti-Individualismus als Problem für mentale Verursachung*.

An dieser Stelle möchte ich nun eine ausführliche Beschreibung meiner Vorgehensweise darbieten. Den Ausgangspunkt für meine Untersuchungen bildet Descartes' Theorie des Geistes. Denn für einen der Väter der Philosophie der Moderne ist die stimmige Erläuterung mentaler Kausalität ein diffiziles Unterfangen. Die Schwierigkeiten, die Descartes bei der Erklärung der kausalen Wechselwirkung zwischen Mentalem und Physikalischem hat, werden gemeinhin als Interaktionsproblem bezeichnet. Genau genommen tritt das, was unter dem Begriff ‚Interaktionsproblem‘ zusammengefasst wird, in unterschiedlichen Formen in Erscheinung.

Die erste Spielart des Interaktionsproblems, die ich in meiner Arbeit analysiere, befasst sich mit der Forderung einer *räumlichen Lokalisierbarkeit von Ursache und Wirkung*, sowie mit der Existenz eines *kausalen Nexus*. Wenn nämlich die räumliche Lokalisierbarkeit von Ursa-

che und Wirkung eine notwendige Bedingung für alle Kausalbeziehungen ist, dann ist mentale Verursachung im Rahmen des interaktionistischen Substanzdualismus a priori offensichtlich unmöglich. Schließlich ist die Nicht-Räumlichkeit das zentrale Charakteristikum des Mentalen in Descartes' Modell. Folglich kann Mentales unter diesen Umständen prinzipiell nicht in Kausalbeziehungen treten. Ein ähnlicher, allerdings etwas abgeschwächter, Einwand ist die Forderung nach einem kausalen Nexus, d.h. einer Schnittstelle, an der Ursache und Wirkung interagieren können. Die Vorstellung eines kausalen Nexus war in der Philosophie zu Zeiten Descartes' weit verbreitet. So wurde die Natur von vielen³⁰ als eine Art gigantisches Uhrwerk betrachtet, in dem alle (kausalen) Vorgänge eindeutig nachvollziehbar sind – analog zu den Prozessen (in) einer Uhr, die in vollem Umfang durch das Ineinandergreifen von Zahnrädern, das Spannen und Lösen von Federn, die Schwingung eines Pendels, etc. verstanden werden können. Diese stark mechanistisch geprägte Auffassung von Kausalität ist auch maßgeblich für das Postulat der Existenz eines kausalen Nexus verantwortlich. Denn dort treibt die Ursache die Wirkung gewissermaßen an und an dieser Stelle findet auch die eigentliche kausale Interaktion statt. Im Rahmen von Kausalbeziehungen innerhalb der Physik ist der kausale Nexus eindeutig räumlicher Natur. Für mentale Kausalität im interaktionistischen Substanzdualismus kann dies allerdings augenscheinlich nicht gelten. Auf die Frage nach der Beschaffenheit des kausalen Nexus im Falle einer kausalen Interaktion zwischen Mentalem und Physikalischem scheint der Substanzdualismus nicht in der Lage zu sein eine befriedigende Antwort geben zu können. Wieder ist es die immanente Nicht-Räumlichkeit des Mentalen, die dem Verständnis mentaler Verursachung zuwiderläuft.

Ein Gegenargument zu diesem Einwand könnte z.B. lauten, dass weder die räumliche Lokalisierbarkeit von Ursache und Wirkung noch das Vorhandensein eines kausalen Nexus notwendige Bedingungen für Verursachungsbeziehungen sind. Denn selbst in der Physik ist es nicht in allen Fällen erforderlich, dass räumlicher Kontakt zwischen Ursache und Wirkung besteht. So ist etwa davon auszugehen, dass die Gravitation eine kausale Kraft ist. Nichtsdestotrotz übt ein Körper auf einen anderen diese Form der Anziehungskraft aus ohne dabei ersichtlichen räumlichen Kontakt zu haben.

Eine zweite Facette des Interaktionsproblems gründet auf den *Erhaltungssätzen* der Physik, hierbei insbesondere auf dem Impuls- und dem Energieerhaltungssatz. Eng mit diesem Punkt assoziiert ist die Forderung nach der *Vollständigkeit* bzw. der *kausalen Geschlossenheit der Physik*. Die Erhaltungssätze der Physik besagen, vereinfacht ausgedrückt, dass eine bestimm-

³⁰ Descartes selbst vertrat genau diesen Standpunkt in Hinblick auf die gesamte physische Welt. Und selbst Tiere waren für ihn reine Automaten. Das Einzige, das nicht unter dieses Diktat der Mechanik unterworfen ist, ist der menschliche Geist.

te Größe innerhalb eines geschlossenen Systems unter Idealbedingungen stets konstant bleibt. Daher ist es beispielsweise fraglich, wie mentale Ursachen physikalische Wirkungen auslösen können. Bekanntermaßen ist der menschliche Körper Teil der physischen Welt. Wenn aber mentale Ursachen kausal verantwortlich für Körperbewegungen sind, müsste etwas von außerhalb des physikalischen Systems entweder einen Impuls oder irgendeine Form von Energie in das System einbringen. Genauso verhält es sich auch in umgekehrter Richtung, d.h. wenn physikalische Ursachen zu mentalen Veränderungen führen. Denn bei dieser Art der kausalen Interaktion müsste ein Teil der Gesamtenergie bzw. des Gesamtimpulses vom Bereich des Mentalen absorbiert werden. Insofern die Erhaltungssätze der Physik gelten, scheint eine kausale Interaktion zwischen Mentalem und Physikalischem fragwürdig. Eine bedeutsame Prämisse für diesen Gedankengang ist die Behauptung der Vollständigkeit bzw. der kausalen Geschlossenheit der Physik. Eine verbreitete Lesart dieses Prinzips besagt, dass jede physikalische Wirkung eine hinreichende physikalische Ursache besitzt. Unter diesen Umständen scheint man – wenn es um das Verursachen von physikalischen Wirkungen geht – vor der Alternative zu stehen entweder das Mentale aus dem kausalen Geschehen gänzlich auszuschließen oder von einer konstanten Überdeterminierung durch mentale und physikalische Ursachen auszugehen.

Gegen diese zweite Abwandlung des Interaktionsproblems ließe sich Folgendes einwenden: nur bestimmte Kausalitätstheorien fordern die Übertragung einer Erhaltungsgröße von der Ursache auf die Wirkung. Lehnt man jedoch diese sogenannten Transfer-Theorien ab, scheint dadurch auch die Verpflichtung zu entfallen, dass bei einer Kausalbeziehung eine konstante Größe übertragen werden muss. Auf diese Weise können die Erhaltungssätze der Physik ihre Allgemeingültigkeit wahren, ohne dabei mentale Ursachen von physikalischen Wirkungen prinzipiell auszuschließen. Im Hinblick auf die kausale Geschlossenheit der Physik wäre es eine Option für interaktionistische Dualisten von einer systematischen Überdeterminierung auszugehen. So könnten beispielsweise gewisse physikalische Wirkungen sowohl eine hinreichende physikalische Ursache als auch eine hinreichende mentale Ursache besitzen.

Eine dritte Variante des Interaktionsproblems ist das sogenannte *Pairing Problem*. Im Kern moniert das Pairing Problem, dass es keine Zuordnungsfunktion gibt, die mentalen Ursachen eineindeutig ihre Wirkungen zuweist. Dazu wird folgendes Gedankenexperiment angeführt: zwei mentale Entitäten M_1 und M_2 üben gleichzeitig einen identischen mentalen Akt m aus, der die Ursache einer Veränderung im Bereich des Physikalischen V_P ist.³¹ Für interaktionisti-

³¹ Das Pairing Problem betrifft nicht nur die kausale Interaktion zwischen Mentalem und Physikalischem, sondern lässt sich auch auf die kausale Wechselwirkung zwischen jeweils zwei mentalen Ursachen und Wirkungen anwenden. Auch hier sei keineswegs klar wie die kausale Paarbildung vonstatten gehen sollte.

sche Substanzdualisten ist es nun unmöglich, so der Vorwurf, zu sagen, ob M_1 s oder M_2 s Vollführen von m die Ursache von V_P ist. Im Gegensatz dazu garantieren bei ähnlichen Fällen im Gebiet der Physik entweder die räumlichen Verhältnisse der Entitäten zueinander oder die Etablierung einer lückenlosen Kausalkette eine eindeutige Zuordnung der kausalen Paare. Auch in diesem Fall ist es die Nicht-Räumlichkeit des Mentalen, die mit dem Verständnis psychophysischer Kausalität konfligiert.

Als Lösung für das Pairing Problems werden beispielsweise reziproke dispositionale Eigenschaften angeführt. So besitzt etwa ein Salzkristall die dispositionale Eigenschaft sich in Wasser auflösen zu können und Wasser hat die dispositionale Eigenschaft Salzkristalle auflösen zu können. Analog verfügt eine Seele über die Eigenschaft mit genau ‚ihrem‘ Körper kausal interagieren zu können (und vice versa). Ein weiterer Lösungsvorschlag gesteht Seelen individualistische Kräfte zu, gemäß derer sie in der Lage sind kausal mit bestimmten Objekten zu interagieren und gleichzeitig nicht mit intrinsisch identischen Entitäten in kausale Wechselwirkung treten. Ein dritter Ansatz führt personengebundene Gesetze ein, die eine konstante Einheit einer bestimmten Seele mit einem bestimmten Körper festsetzen und somit eine genaue kausale Zuordnung garantieren.

Descartes' eigene Lösung des Interaktionsproblems ist in gewisser Weise zweigeteilt. Einerseits lokalisiert er die kausale Interaktion zwischen Körper und Seele an einer Schnittstelle im Körper – nämlich der Zirbeldrüse im Gehirn. An dieser Stelle können mentale und physische Substanz kausal aufeinander wirken. Andererseits legt Descartes auch mehrere nicht weiter definierbare Grundbegriffe des Denkens für seine Theorie des Geistes fest. Wobei der Grundbegriff für die körperliche Substanz die Ausgedehntheit, der Grundbegriff für die mentale Substanz das Denken und der Grundbegriff für beide zusammen die Einheit ist. Eine weitere Lösung für alle Spielarten des Interaktionsproblems wäre es zwischen zwei fundamentalen Arten der Kausalität zu unterscheiden – und zwar der physikalischen Kausalität und der mentalen Kausalität. Beide repräsentieren jeweils einen Grundtypus von Verursachung und sind nicht aufeinander reduzierbar.

Die bekannteste Formulierung des *Problems um die Anomalie des Mentalen* stammt von Donald Davidson. Davidson illustriert diese Schwierigkeit mentale Verursachung stimmig darzustellen anhand eines Trilemmas. D.h. mithilfe dreier Prinzipien, die intuitiv allesamt zur Zustimmung einladen, von denen jedoch immer zwei zusammengenommen scheinbar die Negation des dritten beinhalten. Diese sind:

1. Prinzip der kausalen Interaktion: Manche mentalen Ereignisse interagieren kausal mit physikalischen Ereignissen.
2. Prinzip des nomologischen Charakters von Kausalität: Ereignisse, die als Ursache und Wirkung miteinander verbunden sind, instanziierten immer ein striktes deterministisches Gesetz.
3. Prinzip der Anomalie des Mentalen: Es gibt keine strikten, deterministischen Gesetze, auf deren Basis mentale Ereignisse vorhergesagt oder erklärt werden könnten.

Was die Plausibilität der einzelnen Prinzipien betrifft, so scheint das erste Prinzip über jeden Zweifel erhaben zu sein. Zumindest insofern man an der Existenz mentaler Verursachung festhalten möchte. Nach dem zweiten Prinzip weisen alle Kausalbeziehungen einen nomologischen Charakter auf. Diese Aussage ist stark an das Standard-Erklärungsschema der empirischen Wissenschaften für Kausalzusammenhänge angelehnt. Nach dem Deduktiv-Nomologischen-Modell leitet eine wissenschaftliche Erklärung aus Anfangsbedingungen und einem oder mehreren (Natur-)Gesetz(en) eine Schlussfolgerung ab. Dementsprechend instanziierten Kausalrelationen stets ein striktes Gesetz. Das dritte Prinzip wiederum behauptet, dass es keine strikten deterministischen psychophysischen Gesetze gibt. Zwar mag es durchaus Gesetze wie ‚Wenn jemand Schmerz empfindet, wird er versuchen sich von der vermeintlichen Quelle des Schmerzens zu entfernen‘ geben. Allerdings beinhalten solche Gesetze stets *ceteris paribus*-Klauseln und haben somit nicht den Status eines allgemeingültigen Naturgesetzes. Folglich können die Relata einer echten Kausalbeziehung solche Gesetze nicht instanziierten. Offenbar ist es *prima facie* unmöglich alle drei Prinzipien zusammen kohärent zu vertreten: denn geht man von der Wahrheit von 1. und 2. aus, so ist man allem Anschein nach zur Annahme von strikten psychophysischen Gesetzen verpflichtet. Wer wiederum 1. und 3. vertritt, muss am nomologischen Charakter aller Kausalbeziehungen zweifeln. Die Vereinigung der Prinzipien 2. und 3. schließlich, scheint nahezu legen, dass es keine kausale Interaktion zwischen Mentalem und Physikalischem gibt.

Davidson zufolge besteht der Widerspruch zwischen den drei Prinzipien jedoch nur oberflächlich und löst sich bei einer genaueren Betrachtung auf. Seine kongeniale Lösung für das Anomalie-Problem besteht hauptsächlich in der These, dass jedes mentale Ereignis zugleich ein physikalisches Ereignis sein muss. Dabei versteht Davidson unter ‚Ereignis‘ ein nicht-wiederholbares, zeitlich eindeutig datiertes Einzelding. Erst die Art der Bezugnahme macht ein Ereignis zu einem mentalen oder zu einem physikalischen Ereignis. Der scheinbare Wi-

derspruch zwischen den drei Prinzipien kann folgendermaßen überwunden werden: Identität und Kausalität sind Beziehungen zwischen Ereignissen unabhängig von deren Beschreibung. Da Gesetze sprachliche Konstrukte sind, spielt für sie die Art wie auf Ereignisse referiert wird eine entscheidende Rolle. So ist die Erklärung und Vorhersage von Ereignissen mitunter nur unter einer bestimmten Beschreibung möglich. Genau genommen, besagt das Prinzip 2. lediglich, dass es, wenn zwei Ereignisse in einem Kausalverhältnis zueinander stehen, *eine* Art der Beschreibung dieser beiden Ereignisse gibt, unter der sie ein striktes deterministisches Gesetz instanziierten. Allerdings verpflichtet das Prinzip 2. nicht zur Aussage, dass *jede* wahre singuläre Kausalaussage ein solches Gesetz instanziiert. Wenn nun also ein mentales Ereignis e_1 ein physikalisches Ereignis e_2 verursacht (oder umgekehrt), dann heißt zunächst nur, dass zwei Ereignisse in einem Kausalverhältnis stehen und auf mindestens eines davon kann mit mentalistischem Vokabular Bezug genommen werden. Da Davidson die Welt als im Grunde physikalisch begreift, müssen alle Ereignisse notwendigerweise (auch) eine physikalische Beschreibung besitzen. Deswegen muss für e_1 neben der mentalen Beschreibung (e_1^m) zusätzlich noch eine physikalische (e_1^p) vorhanden sein. Auf diese Weise verschwindet das kontradiktorische Spannungsverhältnis zwischen den drei Prinzipien. Schließlich instanziiert das Kausalverhältnis zwischen e_1 und e_2 ein striktes, deterministisches Gesetz, wenn man sich auf e_1 mittels der physikalischen Beschreibung (e_1^p) bezieht. Wird jedoch auf e_1 qua mentaler Beschreibung (e_1^m) Bezug genommen, so stehen e_1 und e_2 *nicht* in einem nomologischen Verhältnis zueinander und somit gilt auch Prinzip 3. Seine Position nennt Davidson Anomalen Monismus.

Ein Einwand gegenüber Davidsons Konzeption ist, dass e_1 hier e_2 nicht qua Mentalem, sondern qua Physikalischem, verursacht. Auf diese Weise wird die gesamte kausale Arbeit vom Physikalischen verrichtet und das Mentale verkümmert letztendlich zu einem Epiphänomen. Diesen Vorwurf könnte man mit folgendem Beispiel zu untermauern versuchen: Angenommen, man legt einige Birnen auf eine Waageschale und der Zeiger der Waage geht zur Zwei-Pfund Marke. In diesem Fall hat, so könnte man sagen, etwas Grünes und Französisches die Bewegung des Zeigers verursacht. Allerdings existieren keine Gesetze, die das erste Ereignis aufgrund der Grünheit und der französischen Herkunft der Birnen mit dem zweiten Ereignis kausal in Verbindung bringen. Das wiederum ist gleichbedeutend damit, dass die Eigenschaften grün zu sein bzw. aus Frankreich zu kommen des Ursachenereignisses nicht kausal relevant für das Wirkungsereignis waren. Zwar mögen die Eigenschaften durchaus dafür relevant gewesen sein, dass das Ursachenereignis das *Ereignis* war, das es war, aber nicht dafür, dass es die *Ursache* war, die es war. Das lässt den Schluss zu, dass Verursa-

chungsbeziehungen nur aufgrund bestimmter Eigenschaften – eben der kausal relevanten Eigenschaften – einen nomologischen Charakter aufweisen. Substituiert man das zweite Prinzip in Davidsons Trilemma mit diesem Prinzip vom nomologischen Charakter kausal relevanter Eigenschaften, so kann man zu dem Schluss kommen, dass mentale Eigenschaften in Davidsons Theorie nicht kausal relevant sind.

Eine Möglichkeit den Anomalen Monismus gegen den Epiphänomenalismus-Vorwurf zu verteidigen liegt in einer Unterscheidung von zwei Arten kausaler Relevanz. Demnach ist die Aussage ‚Die Eigenschaft F eines Ereignisses c ist kausal relevant für die Eigenschaft G eines Ereignisses e‘ ambivalent. Zum Einen kann es bedeuten die Eigenschaften F und G sind kausal relevant₁ dafür, dass es der Fall ist, dass c e verursacht. Zum anderen kann der Satz aber auch ausdrücken, dass c’s Haben der Eigenschaft F kausal relevant₂ für e’s Besitzen der Eigenschaft G ist. Die erste Aussage gilt also genau dann, wenn c F und e G hat und es ein striktes Gesetz gibt, das es nach sich zieht, dass Fs Gs verursachen. In diesem Sinne sind c’s Haben von F und e’s Besitzen von G dafür verantwortlich, dass c e verursacht. Im Gegensatz dazu ist der zweite Satz genau dann gültig, wenn die Tatsache, dass c F hat, bewirkt, dass e G hat. Diejenigen, die dem Anomalen Monismus entgegenhalten, das Mentale zu einem Epiphänomen zu machen, setzen beide Lesarten von kausaler Relevanz fälschlicherweise gleich. Denn der Anomale Monismus hat zwar zur Folge, dass mentale Eigenschaften kausal irrelevant₁ sind, er impliziert jedoch nicht, dass sie auch kausal irrelevant₂ sind. Denn kausale Irrelevanz₂ ist lediglich eine Form der kontrafaktischen Abhängigkeit zwischen den Eigenschaften F von c und G von e. Dergestalt können mentale Eigenschaften auch durchaus kausal relevant sein. Nun stellt sich die Frage, ob die kausale Irrelevanz₁ mentaler Eigenschaften bereits den Epiphänomenalismusvorwurf gegenüber dem Anomalen Monismus rechtfertigt. Dies scheint unwahrscheinlich zu sein, denn die meisten Eigenschaften würden ihrer kausalen Relevanz beraubt werden, wenn das Fallen unter strikte Gesetze dafür notwendig wäre. Somit kann beim Prinzip von nomologischen Charakter kausal relevanter Eigenschaften auch nur kausale Relevanz₂ gemeint sein. Ansonsten könnte beispielsweise die Eigenschaft ‚ist ein Streichholzreiben‘ nicht kausal relevant für die Eigenschaft ‚ist ein Streichholzentzünden‘ sein. Dies wiederum wäre höchst kontraintuitiv.

Ein sehr ähnlicher Punkt ist die Behauptung das Problem um die Anomalie des Mentalen beruhe auf einer falschen Prämisse. Wenn es nämlich wahr wäre, dass Kausalbeziehungen immer ein striktes Gesetz instanziiieren müssen, dann gäbe es in den Spezialwissenschaften keinerlei Kausalgesetze. Somit wären auch die meisten Eigenschaften, die in den Spezialwissenschaften eine kausale Rolle übernehmen, epiphänomenal. Dies erachtet Fodor als absurd und

zieht den Schluss, dass als eine Mindestvoraussetzung einer Theorie der kausalen Verantwortung von Eigenschaften Folgendes gelten muss: Multipel realisierbare Eigenschaften wie die Eigenschaft eine Tragfläche, ein Segel, ein Berg, eine Annahme, oder ein Verlangen zu sein dürfen nicht von vornherein von der kausalen Verantwortung ausgeschlossen werden. Allerdings gilt es auch zu beachten, dass die Identifikation der kausal verantwortlichen Makroeigenschaft nicht mit der genauen Angabe des implementierenden Mechanismus gleichzusetzen ist. Außerdem unterscheidet sich auch der Wortschatz der beiden Ebenen voneinander. In der Genetik beispielsweise verwendet man für die Gesetze der Vererbung Begriffe wie dominante oder rezessive Merkmale, wohingegen beim implementierenden Mechanismus die Rede von Genen, Chromosomen und DNS ist. Obwohl die Gesetze in den Spezialwissenschaften nicht strikt sind, handelt es sich dabei dennoch um Kausalgesetze. Denn selbst für den Fall, dass Ursachen von Gesetzen abgedeckt werden müssen, die ihre Wirkungen notwendig machen, so müssen diese Gesetze nicht zwangsläufig strikt sein. Schließlich machen Ursachen selbst dann ihre Wirkungen notwendig, wenn sie von eingeschränkt gültigen Gesetzen abgedeckt werden und zugleich deren *ceteris paribus* Klausel erfüllt ist. Insofern existieren auch intentionale Kausalgesetze und bewahren damit mentale Eigenschaften vor dem Stigma des Epiphänomenalismus.

Ein weiteres Argument, das die kausale Wirksamkeit mentaler Eigenschaften in Zweifel zieht ist das sogenannte *Supervenienz-* oder *Exklusionsargument* von Jaegwon Kim. Letztlich versucht das Argument zu zeigen, dass mentale Verursachung im Rahmen des nicht-reduktiven Physikalismus nicht möglich ist. Als Ansatzpunkt dienen drei Prinzipien, denen wohl alle Anhänger des nicht-reduktiven Physikalismus zustimmen würden – nämlich Supervenienz, Nicht-Reduzierbarkeit und Kausale Wirksamkeit. Die Supervenienzthese besagt, dass mentale Eigenschaften stark auf physikalischen Eigenschaften supervenieren. D.h. wenn ein System *s* zum Zeitpunkt *t* eine mentale Eigenschaft *M* instanziiert, dann existiert notwendigerweise auch eine physikalische Eigenschaft *P*, die von *s* zum Zeitpunkt *t* instanziiert wird. Außerdem gilt notwendigerweise, dass wenn etwas *P* zu irgendeinem Zeitpunkt instanziiert, dann instanziiert es zugleich auch *M*. Das Nicht-Reduzierbarkeitsprinzip macht geltend, dass mentale Eigenschaften nicht auf physikalische Eigenschaften reduzierbar sind. Das Prinzip der kausalen Wirksamkeit wiederum schreibt mentalen Eigenschaften kausale Wirksamkeit zu. Zwei weitere – für nicht-reduktive Physikalisten mutmaßlich unstrittige – Leitsätze sind nach Kim noch nötig, um die kausale Unwirksamkeit mentaler Eigenschaften zu beweisen. Diese Maximen sind erstens die kausale Geschlossenheit der Physik und zweitens das Prinzip der kau-

salen Exklusion. Das Geschlossenheitsprinzip geht davon aus, dass, wenn ein physikalisches Ereignis eine Ursache zum Zeitpunkt t hat, es an t eine physikalische Ursache hat. Das Exklusionsprinzip behauptet, wenn ein Ereignis eine hinreichende Ursache hat, dann hat es nicht zugleich eine andere, von der ersten verschiedene, hinreichende Ursache; es sei denn es handelt sich um einen Fall genuiner kausaler Überdeterminierung.

Der erste Schritt des Exklusionsarguments geht von einem ordinären Fall von Mental-zu-mental-Verursachung aus. Die (Instanziierung einer) mentale(n) Eigenschaft M verursacht die (Instanziierung einer) andere(n) mentale(n) Eigenschaft³² M^* . Der zweite Schritt ist schlichtweg die Anwendung der Supervenienzthese auf dieses Beispiel. Für irgendeine physikalische Eigenschaft P^* muss gelten: M^* hat P^* als Supervenienzbasis. So entsteht ein Spannungsverhältnis zwischen M , der vermeintlichen Ursache von M^* , und seiner Supervenienzbasis P^* in Hinblick auf die Frage ‚Warum wird M^* instanziiert?‘. Das Supervenienzprinzip legt jedoch fest, dass M^* notwendigerweise stattfinden muss, wenn P^* instanziiert wird – also selbst dann, wenn vorher M nicht eingetreten wäre. Zudem ist das alleinige Eintreten von P^* schon hinreichend für die Instanziierung von M^* . Daher scheint die einzige Möglichkeit zur Rettung der kausalen Wirksamkeit von M im dritten Schritt des Arguments zu liegen. Der besteht in der Behauptung M habe M^* verursacht, indem es dessen Supervenienzbasis P^* verursacht hat. Damit ist das erste Stadium des Arguments abgeschlossen. Dieses zeigt nach Kim, dass Mental-zu-mental-Verursachung im nicht-reduktiven Physikalismus notgedrungen Mental-zu-physikalisch-Verursachung beinhaltet.

Für den Abschluss des Arguments gibt es zwei unterschiedliche Varianten. Bei der ersten Vervollständigung ist der vierte Schritt des Exklusionsarguments die Anwendung der Supervenienzthese auf M . So muss auch M eine Supervenienzbasis besitzen, und zwar P . Unmittelbar daraus leitet sich als fünfte Prämisse ab: M verursacht P^* und P verursacht P^* . Der sechste Schritt des Arguments ist die Applikation der Nicht-Reduzierbarkeit auf das Beispiel – demnach sind M und P notwendigerweise nicht identisch. Aus den Annahmen fünf und sechs ergibt sich ein Wettbewerb um den Status als Ursache von M^* . An dieser Stelle kommt das Exklusionsprinzip zur Geltung, demzufolge es nicht zwei hinreichende Ursachen für ein und dasselbe Ereignis geben kann (es sei denn es handelt sich um einen Fall genuiner Überdeterminierung). Dies ist auch der siebte Schritt des Arguments: P^* ist nicht kausal überdeterminiert durch P und M . D.h. eine der beiden angeblichen Ursachen von P^* muss fallengelassen werden. Als Entscheidungsgrundlage dafür welche verworfen werden muss dient das Ge-

³² Streng genommen müsste es heißen ‚Die Instanziierung der mentalen Eigenschaft M verursacht die Instanziierung einer anderen mentalen Eigenschaft M^* ‘. Im weiteren Textverlauf werde ich jedoch genau wie Kim auf die verkürzte Redeweise von ‚ M verursacht M^* ‘ zurückgreifen.

geschlossenheitsprinzip. Infolgedessen kann es nur die mentale Eigenschaft M treffen. Daher kommt das Exklusionsproblem im achten Schritt zur Konklusion: die vermeintliche mentale Ursache M wird von der physikalischen Ursache P ausgeschlossen, somit ist P die eigentliche Ursache von P* und nicht M.

Die zweite mögliche Vollendung des Arguments hat die Form einer *reductio ad absurdum*. Bei dieser Variante des Exklusionsarguments ist vierte Prämisse die Aussage ‚M ist die Ursache von P‘. Der fünfte Schritt des Arguments ergibt sich aus dem Geschlossenheitsprinzip. Da es sich bei P* um die Instanziierung einer physikalischen Eigenschaft handelt, benötigt P* auch eine physikalische Ursache P, die zum selben Zeitpunkt eintritt wie M. Akzeptiert man zugleich die Nicht-Reduzierbarkeit mentaler Eigenschaften, so erhält man als sechsten Schritt des Arguments die Aussage, dass M nicht mit P identisch ist. Dies hat wiederum die siebte Prämisse zur Folge, die besagt, dass P* zwei unterschiedliche Ursachen hat – nämlich M und P – und dies kein Fall genuiner Überdeterminierung ist. Nach dem Exklusionsprinzip kann der achte Schritt des Arguments nur lauten: entweder muss M oder P als Ursache von P* fallengelassen werden. Erneut bringt das Geschlossenheitsprinzip die Entscheidung und legt die Konklusion des Exklusionsarguments fest. Diese besagt, dass M als Ursache* von P ausgeschlossen werden muss und somit nur P erhalten bleibt. Damit ist die *reductio* vollzogen und laut Kim bewiesen, dass physikalische Eigenschaften die kausale Wirksamkeit mentaler Eigenschaften im nicht-reduktiven Physikalismus verhindern.

Ein Einspruch gegen das Exklusionsargument setzt am Herzstück des Arguments – dem Exklusionsprinzip – an. Gemäß des Exklusionsprinzips schließen sich nämlich nur zwei voneinander unabhängige hinreichende Ursachen gegenseitig aus. Im Falle von Geist-Körper-Supervenienz ist diese Unabhängigkeit aber gerade nicht gegeben. Außerdem handelt es sich z.B. bei M und P* um zwei unterschiedliche Arten der Erklärung für M*. Während M eine diachronische, kausale Erklärung für M* ist, liefert P* eine synchrone, nicht-kausale Erklärung für das Eintreten von M*. Analog lässt sich beispielsweise die Frage ‚Warum hat Max Masern?‘ auf zwei unterschiedliche Weisen korrekt beantworten. Erstens könnte man sagen, dass Max Masern hat, weil er sich mit einem bestimmten Virus infiziert hat. Zweitens könnte man auch sagen, dass Max Masern hat, weil er bestimmte physikalische Basiseigenschaften instanziiert, auf denen ‚Masern haben‘ superveniert (i.e. Max zeigt die typischen Symptome einer Ansteckung mit Masern). Dies sind beides legitime Erklärungen für Max’ Masern – allerdings stehen sie nicht in einem Konkurrenzverhältnis zueinander.

Ein weiterer Einwand richtet sich direkt gegen das Exklusionsprinzip. Denkt man dieses nämlich konsequent zu Ende, so besteht die Gefahr, dass die kausalen Kräfte auf die fundamental-

ste Ebene der Physik versickern. D.h. strenggenommen gibt es keine mentale, keine physiologische, keine biologische, keine molekulare,... Verursachung – Kausalität spielt sich ausschließlich auf der Ebene der Elementarteilchen ab. Allerdings ist es in der Wissenschaftsgemeinschaft strittig, ob es so etwas wie eine Fundamentalebene in der Physik überhaupt gibt. Wäre nun das Exklusionsprinzip wahr und gäbe es gleichzeitig kein Basislevel der Physik, würden kausale Kräfte ins Bodenlose versickern. In anderen Worten, es gäbe keine Verursachung mehr. Dies ist allerdings höchst kontraintuitiv und widerspricht der gängigen explanatorischen Praxis, die für die Vielzahl der Fälle sehr überzeugend ist.

Ein weiterer Kritikpunkt am Exklusionsargument greift an Kims Verständnis der Instanziierung von Eigenschaften an. Nach Kim können nicht-reduktive Physikalisten nämlich nicht behaupten M und P seien identisch. Das ist jedoch falsch, denn bei M und P handelt es sich nicht um Eigenschaften, sondern um Instanziierungen von Eigenschaften. So könnte ein Verfechter des nicht-reduktiven Physikalismus durchaus den Standpunkt vertreten die Instanziierung von M sei mit der Instanziierung von P identisch, aber die Eigenschaft M-zu-sein und die Eigenschaft P-zu-sein seien jedoch verschieden. Allerdings scheint diese Unterscheidung nach Kims Konzeption der Instanziierung von Eigenschaften nicht möglich zu sein. Es hat nämlich den Anschein als vollziehe sich für Kim die Instanziierung von Eigenschaften analog zu seinem Ereignisbegriff. D.h. instanziiert etwa ein Individuum x eine Eigenschaft F zum Zeitpunkt t, so ergibt sich daraus folgende komplexe, zusammengesetzte Entität $[x, F, t]$. Nach dieser Vorstellung sind die Instanziierungen von zwei Eigenschaften F und G identisch, wenn die Eigenschaften identisch sind und zur selben Zeit vom gleichen Träger instanziiert werden: $[x, F, t] = [x, G, t] \rightarrow F = G$. Damit fällt aber die Unterscheidung zwischen Type- und Token-Identität zusammen. Diese Unterscheidung ist zentral für den nicht-reduktiven Physikalismus. Ein Argument, das die Ablehnung des nicht-reduktiven Physikalismus bereits implizit in seiner Konzeption der Eigenschaftsinstanziierung trägt, kann als zirkulär zurückgewiesen werden.

Eine letzte Beanstandung bezieht sich auf die Überdeterminierungsklausel im Exklusionsprinzip, der zufolge ein Ereignis nicht zwei hinreichende Ursachen haben kann, es sei denn es handelt sich um einen Fall genuiner kausaler Überdeterminierung. Allgemein stellt sich die Frage, ob die überdeterminierenden Ursachen voneinander unabhängig sein müssen oder nicht. Falls dies nämlich nicht der Fall sein muss, könnte bei M und P ein Fall genuiner kausaler Überdetermination vorliegen. Somit wären M und P nicht vom Exklusionsprinzip betroffen. Falls die Unabhängigkeit der überdeterminierenden Ursachen jedoch ein integraler Bestandteil des Konzepts genuiner kausaler Überdeterminierung ist, bleibt offen, ob M und P aufgrund

ihres Abhängigkeitsverhältnisses tatsächlich jeweils als hinreichende Ursache für M^* betrachtet werden können.

Neben den gerade genannten Beanstandungen des Exklusionsarguments gibt es auch zahlreiche Lösungsvorschläge für die aufgeworfene Problematik. Eine Lösungsmöglichkeit liegt in der Auffassung mentale Eigenschaften seien die Determinablen physikalischer Eigenschaften, ihrer Determinaten und stehen deshalb in keinem Konkurrenzverhältnis zueinander. Ein Beispiel für eine Determinationsbeziehung ist das Verhältnis von scharlachrot zu rot zu farbig. Während scharlachrot eine Determinate der Determinable rot (oder auch der Determinable farbig) ist, verhält sich rot zu farbig wie Determinate zu Determinable. Was nun die kausale Wirksamkeit betrifft, so sind Determinaten und Determinablen keine Wettbewerber im Hinblick auf ihre Kausalkräfte. Zur Illustration dieses Gedanken betrachte man folgendes Szenario: eine Taube wird so konditioniert, dass sie bei der Präsentation von mehreren farbigen Gegenständen mit ihrem Schnabel immer auf den roten Gegenstand pickt. Nun wird dem Vogel ein scharlachroter Gegenstand vorgelegt, auf den sie – wie erwartet – pickt. Eine Übertragung des gedanklichen roten Fadens des Exklusionsarguments auf dieses Beispiel könnte wie folgt aussehen: zunächst erschien es so als wäre die Eigenschaft rot-zu-sein kausal relevant für das Picken gewesen. Bei näherer Betrachtung stellt sich jedoch heraus, dass es die Eigenschaft scharlachrot-zu-sein war, die in Wirklichkeit kausale Relevanz besaß. Die Eigenschaft scharlachrot-zu-sein schließt die Eigenschaft rot-zu-sein somit vom kausalen Geschehen aus. Diese Beweisführung ist jedoch nicht überzeugend, weil die Eigenschaften rot-zu-sein und scharlachrot-zu-sein nicht darum wetteifern Ursache des Pickens zu sein. Genausowenig stehen mentale und physikalische Eigenschaften im Wettbewerb.

Eine weitere Antwort auf das Exklusionsproblem bezieht sich nicht speziell auf mentale und physikalische Eigenschaften, sondern befasst sich ganz allgemein mit der Frage nach dem Beitrag, den höherstufige Eigenschaften grundsätzlich in Kausalbeziehungen leisten können. Zu diesem Zweck wird eine genaue Unterscheidung zwischen ‚kausaler Wirksamkeit‘ und ‚kausaler Relevanz‘ eingeführt. Während kausale Wirksamkeit den Eigenschaften vorbehalten bleibt, die tatsächlich in Ursache-Wirkungs-Verhältnissen stehen, können höherstufige, abstrakte Eigenschaften durchaus kausale Relevanz besitzen. Dies erreichen sie indem ihre Instanziierung sicherstellt, dass eine kausal wirksame Eigenschaft eintritt. Obwohl die höherstufige Eigenschaft nicht kausal wirksam ist, also keinerlei tatsächliche kausale Arbeit leistet, kann sie nichtsdestotrotz kausal relevant sein, indem sie die Instanziierung einer kausal wirksamen Eigenschaft gewährleistet. Dabei spielt es keine Rolle welche kausal wirksame Eigenschaft letzten Endes de facto realisiert wird. Ähnlich einer Computersoftware ‚programmier-

ren' höherstufige Eigenschaften das Eintreten der niedrigstufigen Eigenschaften, die schlussendlich die kausale Arbeit verrichten. Wird ein Computerprogramm ausgeführt, so garantiert dies in der Regel, dass gewisse Dinge geschehen werden; die Vorgänge, die diese Dinge jedoch faktisch hervorbringen, liegen auf einer tieferen, mechanischen Ebene. Kurzum, die höherstufigen kausal relevanten Eigenschaften liefern Programmerklärungen, wohingegen die niedrigstufigen kausal wirksamen Eigenschaften Prozesserkklärungen abgeben. Das Gros der Kausalerklärungen in Wissenschaft und Alltag besteht gewiss aus Programmerklärungen, weil oft die realen kausal wirksamen Eigenschaften gar nicht bekannt sind. Zudem haben Programmerklärungen eine eigene Existenzberechtigung, zumal sie manchmal Informationen enthalten, die in der zugehörigen Prozesserkklärung nicht vorhanden sind. Man könnte sagen, dass Prozesserkklärungen verraten wie die tatsächliche kausale Geschichte verlaufen ist, während Programmerklärungen angeben wie die kausale Geschichte hätte verlaufen können. Abschließend dazu ein Beispiel: Die Eigenschaft eines Uranklumpens, dass in diesem Moment ganz bestimmte Atomkerne – und zwar genau jene – zerfallen, verursacht dessen radioaktive Strahlung und ist somit kausal wirksam. Die höherstufige, abstrakte Eigenschaft, dass eine kritische Anzahl von (beliebigen) Atomkernen im Klumpen zerfallen muss, damit er Radioaktivität emittiert, ist kausal relevant. Wird sie nämlich instanziiert, tritt zweifellos eine kausal wirksame Eigenschaft ein. Ob es nun allerdings genau jene Atomkerne sind, die zerfallen oder ganz andere, ist für die (auf der) Makroebene (vom Uranklumpen ausgehende Strahlung) letztlich nicht bedeutsam.

Die nächste Strategie zur Beseitigung des Exklusionsproblems schlägt eine Unterscheidung zweier Gebrauchsweisen des Begriffs ‚Eigenschaft‘ vor. Einerseits wird das Wort ‚Eigenschaft‘ verwendet wie Universalie oder Klasse. In diesem Sinne wäre die Eigenschaft F die allgemeine F-heit aller Dinge, die die Eigenschaft F besitzen oder in anderen Worten die Klasse aller Fs. Andererseits kann der Terminus ‚Eigenschaft‘ auch als abstraktes Partikular oder Trope aufgefasst werden. Liegt diese Lesart zugrunde, dann sind Eigenschaften, in dem Moment, in dem sie instanziiert werden zur Gänze in dem Objekt, das sie gerade instanziiert und können simultan nicht in einem anderen Objekt zugegen sein. Folgendes Beispiel veranschaulicht diese Unterscheidung: zwei reife Bananen haben die Eigenschaft gelb zu sein. Ist die Farbe der einen Banane numerisch die gleiche wie die Farbe der anderen Banane? Versteht man ‚Eigenschaft‘ hier als Universalie, so lautet die Antwort ‚Ja‘, wird der Begriff hingegen als Trope verstanden, muss die Frage negiert werden. Mit dieser Differenzierung kann die vom Exklusionsargument zur Sprache gebrachte Problematik überwunden werden. Begreift man nämlich mentale Eigenschaften als kausal relevant, so ist hier unter Eigenschaft

‚Trope‘ gemeint. Da mentale Tropen allerdings mit physikalischen Tropen identisch sein können, wird das Geschlossenheitsprinzip auf diese Weise nicht verletzt. Im Hinblick auf die Nicht-Reduzierbarkeit mentaler Eigenschaften ist unter dem Begriff hinwieder Universalie zu verstehen.

In einer genauen Analyse der Quausalitätsrelation liegt eine weitere Möglichkeit das Exklusionsproblem zu umgehen. Quausalität ist eine vierstellige Relation der Form ‚c qua F verursacht e qua G‘, wobei c und e Ereignisse, F und G Eigenschaften repräsentieren. Für Quausalität muss gelten, dass c e verursacht. Außerdem muss c F und e G instanziierten. Ferner müssen F und G logisch und metaphysisch unabhängig voneinander sein. Überdies darf die kausale Transaktion zwischen c und e keine Überdeterminierung, Präemption, o. Ä. beinhalten. Schließlich muss noch folgende Relevanzbedingung gelten: die Tatsache, dass c* (ein c angemessen ähnliches Ereignis in einer möglichen Welt w) F instanziiert, in allen angemessen ähnlichen möglichen Welten explanatorisch relevant dafür ist, dass ein e* (ein e angemessen ähnliches Ereignis in einer möglichen Welt w) eintritt und G instanziiert. Angenommen Fred verspürt ein Verlangen nach einem kühlen Bier und dieser Wunsch verursacht eine Handlung – und zwar, dass Fred zum Kühlschrank geht. In diesem Fall besitzt die mentale Ursache c die Eigenschaft ein Verlangen nach einem Bier zu sein und verursacht ein Verhalten e mit der Eigenschaft ein Gang zum Kühlschrank zu sein. Es ist anzunehmen, dass es angemessen ähnliche mögliche Welten gibt, in denen Fred z.B. sein Begehren nach einem Bier in Situationen erlebt, die sich von der aktuellen Welt mäßig unterscheiden (etwa insofern als sich Fred in einem anderen Teil seines Hauses befindet als er seinen Bier-Durst empfindet). Auch in diesen Szenarien würde sein Wunsch nach einem Bier ein Verhalten verursachen, dass sich als Gang zum Kühlschrank qualifiziert. Die entscheidende Frage in diesem Zusammenhang ist, ob es eine angemessen ähnliche mögliche Welt gibt, in der ein Ereignis eintritt, dass Freds Verlangen nach einem Bier angemessen ähnlich ist und das zugleich *nicht* die Eigenschaft hat ein Wunsch nach einem Bier zu sein. Dies erscheint jedoch äußerst kontraintuitiv. Demnach scheint es möglich zu sein, dass mentale Eigenschaften qua Mentalem in Kausalbeziehungen treten – im Gegensatz zum Kerngedanken des Exklusionsarguments.

Ein letzter Themenkomplex, der eine verständliche Darstellung mentaler Verursachung erheblich erschwert, beruht auf dem semantischen *Externalismus* und dem *Anti-Individualismus*. Im Wesentlichen gehen diese beide Ansätze davon aus, dass der Gehalt intentionaler mentaler Zustände extrinsisch individuiert wird. D.h. der Gehalt dieser mentalen Zustände ist abhängig von Faktoren, die *außerhalb* derjenigen Person liegen, die sich gerade in diesem mentalen

Zustand befindet. So können beispielsweise kausale, historische oder soziale Beziehungen, die diese Person zu ihrer Umwelt unterhält, eine entscheidende Rolle spielen. Insgesamt bleibt festzuhalten, dass nicht ausschließlich die intrinsischen Eigenschaften den intentionalen Gehalt mentaler Zustände festlegen. Im Gegensatz dazu scheinen für Kausalbeziehungen allein intrinsische Eigenschaften ausschlaggebend zu sein.

Folgendes Bild konkretisiert diese Schwierigkeit auf eine sehr anschauliche Weise: der menschliche Geist verhält sich zu intentionalen Handlungen wie Geld zu Verkaufsautomaten. Diese Analogie trägt deshalb sehr gut, weil die Individuierungsbedingungen von Geld und von Überzeugungen sehr ähnlich sind. So ist der Geldwert von Münzen z.B. nicht allein den intrinsischen Eigenschaften wie Form, Größe und Dichte zu verdanken, sondern hängt wesentlich von extrinsischen oder relationalen Eigenschaften ab. Um tatsächlich Geldwert zu besitzen, muss eine Münze die richtige Geschichte haben, d.h. sie muss in einer staatlichen Münzprägestalt hergestellt worden sein und nicht in der Garage eines Fälschers. Genauso wenig ist es der intrinsische neurobiologische Charakter einer Überzeugung, der eine Überzeugung zu einer Überzeugung macht. Denn auch in diesem Fall spielen extrinsische Eigenschaften eine ausschlaggebende Rolle. Wie z.B. die Eigenschaft eine bestimmte Bedeutung bzw. einen bestimmten Gehalt zu haben. Nimmt man diese Analogie ernst, dann sind Überzeugungen qua Überzeugung für die Verursachung menschlichen Verhaltens in gleicher Weise bedeutungslos wie es Geld qua monetären Wert für das Verhalten eines Verkaufsautomaten ist. In den Transaktionen mit einem Verkaufsautomaten erklären letzten Endes nur die intrinsischen Merkmale der Münze, warum der Einwurf eines Geldstücks die Ausgabe der Ware verursacht hat. Hätte man nämlich eine gefälschte Münze, die in Form, Größe und Dichte einer echten Münze hinreichend ähnlich ist, in den Automaten eingeworfen, so hätte auch diese den Auswurf des gewünschten Artikels verursacht. Gleichmaßen liefern allein die intrinsischen Charakteristika einer Überzeugung – i.e. ihre neurobiologischen Attribute – und nicht ihr Inhalt eine kausale Erklärung für das Verhalten ihres Trägers. Allem Anschein nach ist das, wovon man überzeugt ist, kausal belanglos für das, was man tut.

Eine Lösung für das Externalismus-Problem liegt im Einführen der Unterscheidung zwischen weitem und engem Gehalt repräsentationaler Zustände. Diese Antwort auf das Externalismus-Problem rekurriert auf ein klassisches Gedankenexperiment zur Untermauerung des Externalismus – das Zwillingerden-Szenario. Die Zwillingerde (oder kurz Zwerde) ist für einen Beobachter ununterscheidbar von der Erde; sie ist allem Anschein nach eine detailgetreue Kopie des blauen Planeten mitsamt all seiner Bewohner. Allerdings unterscheidet sie sich insofern von unserer Welt, als dass Wasser dort nicht die chemische Struktur H_2O aufweist,

sondern einen anderen molekularen Aufbau besitzt – nämlich XYZ. Somit referiert ein Sprecher auf der Zwerde mit der Aussage ‚Das ist Wasser‘ auf etwas anderes als sein irdisches Pendant. Obwohl sich beide in exakt denselben psychologischen und physiologischen Zuständen befinden, unterscheidet sich doch die Bedeutung ihrer Aussagen. In anderen Worten, der Inhalt ihrer Überzeugungen ist ein anderer. An dieser Stelle knüpft die Begriffsdifferenzierung zwischen weitem und engem Gehalt an. Die Kernaussage dieser Strategie ist, dass sich die Aussagen von Erdling und Zwerdling in ihren weiten Gehalten unterscheiden, aber dennoch denselben engen Gehalt haben. Der enge Gehalt wird nämlich vollständig durch die intrinsischen Eigenschaften des Trägers festgelegt. Da diese identisch sind, muss auch der enge Gehalt der mentalen Zustände gleich sein. Der gedankliche Hintergrund dieser Argumentation ist, dass sich die kausale Kraft der mentalen Zustände der beiden Sprecher *nicht* unterscheidet, obwohl sich der eine auf Wasser und der andere auf Zwasser bezieht. Dies liegt daran, dass ein *begrifflich notwendiger* Zusammenhang zwischen Wasser und mentalen Zuständen wie z.B. Überzeugungen, die sich inhaltlich auf Wasser beziehen, besteht. Vor dem Hintergrund des engen Gehalts mentaler Zustände verschwindet auch das Externalismus-Problem.

Ein weiterer Hinweis auf eine Lösung des Externalismus-Problems ist bereits in der Münzautomaten-Analogie enthalten und legt den Grundstein für Fred Dretskes Dual-Explanandum Theorie. Zwar ist es tatsächlich so, dass die intrinsischen Eigenschaften der eingeworfenen Geldstücke die Ausgabe von Verkaufsware verursachen. Allerdings ist es das vitale Interesse von Zentralbanken und Regierungen, dass der Geldwert auf den intrinsischen Eigenschaften superveniert. Zu diesem Zweck versehen sie etwa echtes Geld mit Sicherheitsmerkmalen, die einer Fälschung vorbeugen sollen. Leisten die Behörden dabei gute Arbeit, dann wird der Geldwert notwendigerweise von den intrinsischen Eigenschaften realisiert. Aufgrund dieser starken Korrelation wird der monetäre Wert in geschäftlichen Transaktionen häufig genutzt um die Wirkungen von Geld vorherzusagen und zu erklären. So könnte man beispielsweise die Frage ‚Warum hat mir die Kassiererin 8€ herausgegeben?‘ beantworten, indem man darauf verweist, dass das Mittagessen 12€ gekostet hat und ich ihr einen 20€-Schein ausgehändigt habe. Sicherlich ist diese Art der Erklärung alltäglich und entspricht vollkommen dem gesunden Menschenverstand. Dennoch stellt sich die Frage, ob es wirklich korrekt ist vom Geldwert des Scheines als einer kausal relevanten Eigenschaft auszugehen? Zunächst gilt es zwischen der kausalen Wirksamkeit des Objekts (20€-Schein) und der kausalen Relevanz einer seiner Eigenschaften (20€ wert zu sein) zu unterscheiden. Darüber, dass mein Aushändigen der 20€-Note die Rückgabe des Wechselgelds in Höhe von 8€ *verursacht* hat, besteht

kein Zweifel. Ob jedoch die Eigenschaft 20€ wert zu sein auch *erklärt*, dass ich 8€ Rückgeld erhalten habe ist eine andere Frage. Denn hätte ich der Dame an der Kasse einen perfekt gefälschten 20€-Schein überreicht, dann hätte sie mir gewiss auch 8€ herausgegeben. Somit ist die eigentliche Erklärung für die Herausgabe der 8€ Wechselgeld, dass ich der Kassiererin einen Stück Papier gegeben habe, von dem sie glaubte es sei ein echter Geldschein. Die kausal wirksamen Eigenschaften sind die intrinsischen Eigenschaften der 20€-Note, auf denen ihr Geldwert superveniert. Diese werden für gewöhnlich auch dazu genutzt, um zu bestimmen, ob es sich bei einem Papierstück um eine original Banknote handelt oder nicht. Dieses Beispiel verdeutlicht, dass bei der Erklärung von Kausalzusammenhängen häufig auf extrinsische Eigenschaften zurückgegriffen wird. Nun stellt sich die Frage warum so eine Diskrepanz zwischen der explanatorischen Praxis und den tatsächlichen Kausalverhältnissen besteht. Diese Abweichung hat in erster Linie pragmatische Gründe: oft sind die tatsächlich kausal wirksamen intrinsischen Eigenschaften nicht bekannt. Außerdem können die extrinsischen Eigenschaften, weil sie auf den intrinsischen Eigenschaften supervenieren, durch diese multipel realisiert werden. So gibt es beispielsweise nicht nur eine einzige Münze, die die Eigenschaft 1€ wert zu sein besitzt, sondern sehr viele. Ferner kann der Wert von 1€ auch durch mehrere verschiedene Kombinationen aus Cent-Münzen zustande kommen. In toto ist also die Erklärung kausaler Zusammenhänge anhand extrinsischer Eigenschaften einfacher und eleganter als mittels intrinsischer Eigenschaften. Wenn Überzeugungen mit bestimmten Hirnzuständen identisch sind, dann können deren Eigenschaften in intrinsische (i.e. neurobiologische) und extrinsische (den Inhalt der Überzeugung festlegende) unterschieden werden. Obwohl nun die Inhalte von Überzeugungen oft zur kausalen Erklärung von Verhalten herangezogen werden, sind sie – genau wie der monetäre Wert von Geld – eigentlich nicht ins kausale Geschehen involviert. Bei dieser Betrachtungsweise wird allerdings zumeist ein wichtiger Aspekt übersehen. Im Fall von Verkaufsautomaten beispielsweise sollte man erwägen, dass sie gemäß ihres Bauplans auf bestimmte Eigenschaften reagieren. Diese Form der Konstruktion besteht aber gerade deshalb fort, *weil* der Geldwert V auf diesen intrinsischen Eigenschaften S superveniert. Nimmt man nun an, dass es Kriminellen gelänge große Mengen an Falschgeld mit den Eigenschaften S in Umlauf zu bringen, dann würde man von der Herstellerseite und sicherlich auch von Seiten der Notenbank reagieren. Der Geldwert V würde dann nicht mehr auf S-Objekten supervenieren, sondern würde nun, z.B. durch die Einführung neuer Sicherheitsmerkmal, auf den intrinsischen Eigenschaften T supervenieren. Dementsprechend würden die Münzautomatenfabrikanten ihre Geräte so modifizieren, dass die Maschinen zukünftig nur noch auf Objekte mit den Eigenschaften T Artikel ausgeben. Auf diese Weise verlöre

auch S seine kausale Kraft. Kurzum, Verkaufsautomaten sprechen nur deshalb auf Gegenstände mit den Eigenschaften S an, weil V auf S superveniert, endet die Supervenienz von V auf S, dann können S-Objekte auch nicht mehr die Ausgabe von Ware verursachen. Daraus lässt sich zwar nicht schließen, dass der Wert des Geldes das Auswerfen der Produkte erklärt, aber die Frage ‚Warum gibt der Automat seine Ware aus?‘ kann von untergeordnetem Interesse sein. Vielmehr könnte man hinterfragen ‚Warum verursachen Gegenstände mit diesen und jenen Eigenschaften die Warenausgabe?‘ oder noch abstrakter ‚Weshalb verursachen Objekte dieser Art Wirkungen jener Art?‘. Zusammengefasst erklärt der Geldwert von Münzen also nicht warum Waren aus Verkaufsautomaten ausgegeben werden, aber er erklärt warum Geldstücke den Auswurf von Ware verursachen. In Anwendung auf Fälle mentaler Verursachung bedeutet dies, dass extrinsische Eigenschaften, wie z.B. der Inhalt von Überzeugungen, erklären warum Ereignisse dieser Art Wirkungen jener Art haben. Allerdings ist die innere kausale Struktur von Personen nicht wie im Falle von Automaten der Planung von Ingenieuren geschuldet, sondern allem voran ein Produkt von Evolution und Lernen. Ein Beispiel aus dem Tierreich illustriert dies anschaulich: ein bestimmtes neurologisches Ereignis im visuellen Kortex eines Huhnes wird vom Schatten eines überfliegenden Falken hervorgerufen und verursacht, dass das Huhn niederkauert und sich versteckt. Der Grund dafür, dass Ereignisse dieser Art (bestimmte Aktivierungsmuster im Gehirn von Hühner) Ereignisse jener Art (Ducken und Verstecken) verursachen, liegt in einer wichtigen extrinsischen Eigenschaft des Ursachenereignisses. Der nämlich, dass Ereignisse dieses Typs in der Regel von überfliegenden Falken verursacht werden. Nunmehr ist kein großer argumentativer Schritt mehr nötig, um zu zeigen, dass die extrinsischen Eigenschaften der inneren Ursachen Verhalten kausal erklären und nicht deren intrinsische Gegenstücke. Das ‚missing link‘ hierfür ist eine entsprechende Interpretation des Begriffs ‚Verhalten‘. Darunter sind nämlich nicht die reinen Körperbewegungen zu verstehen, die von inneren Ereignissen verursacht werden, sondern vielmehr ist damit die Verursachung der Körperbewegungen durch innere Prozesse gemeint. Alles in allem erklären also Überzeugungen Verhalten kausal, wohingegen die reinen Körperbewegungen von Hirnzuständen erklärt werden. Insofern ist auch die häufig gezogene Schlussfolgerung aus der Münzautomaten-Analogie, die besagt das Mentale besitze letzten Endes nur epiphänomenalen Charakter, falsch.

Eine weitere grundlegende Unterscheidung für Dretskes Dual-Explanandum Theorie ist die Differenzierung zwischen auslösenden und strukturierenden Ursachen. Hierbei liefern erstere für gewöhnlich eine biologische und letztere eine psychologische Erklärung von Verhalten. Die Verschiedenartigkeit beider Ursachentypen konkretisiert sich an folgenden zwei Beispielen:

len: Drückt ein Computernutzer eine bestimmte Taste und bewegt sich dann der Cursor auf dem Bildschirm, dann ist der Tastendruck die auslösende Ursache der Cursorbewegung. Im Gegensatz dazu setzt sich die strukturierende Ursache für die Cursorbewegung aus den Hard- und Softwarebedingungen zusammen, die bestehen müssen, dass ein Tastendruck dieser Art eine Cursorbewegung jener Art hervorruft. Im zweiten Beispiel platziert ein Terrorist eine Bombe im Auto eines Generals. Als dieser nach einigen Tagen in seinen PKW steigt und den Zündschlüssel umdreht explodiert das Fahrzeug. In diesem Fall ist die Drehung des Zündschlüssels die auslösende Ursache für den Tod des Generals. Die strukturierende Ursache wiederum ist in der hinterhältigen Tat des Attentäters zu sehen. Streng genommen könnte man an dieser Stelle monieren, dass strukturierende Ursachen im eigentlichen Wortsinn gar keine Ursachen sind. Schließlich hängt ihr Dasein als strukturierende Ursachen von späteren Ereignissen ab. Drückt etwa nie jemand die Taste am Computer oder dreht keiner jemals den Zündschlüssel im mit Sprengstoff beladenen Auto, dann sind weder die Computerverdrahtung und –programmierung noch die Platzierung der Autobombe strukturierende Ursachen. Außerdem können strukturierende Ursachen in einer eins-zu-viele-Relation zu ihren Wirkungen stehen, d.h. ein und dieselbe strukturierende Ursache kann viele Wirkungen derselben Art verursachen. In obigem Computer-Beispiel etwa sorgt die Programmierung und Verkabelung des Rechners dafür, dass, wenn eine bestimmte Taste gedrückt wird, jedes Mal aufs Neue eine Cursorbewegung erfolgt. Im Prinzip könnte man auch behaupten, dass es gar keine strukturierenden Ursachen gibt. Denn die strukturierende Ursache S eines bestimmten Ereignisses E ist eigentlich die auslösende Ursache der mehr oder weniger stabilen Hintergrundbedingungen B, in deren Rahmen auslösende Ursachen des Typs T Ereignisse wie E verursachen. In anderen Worten ist S die Ursache von T's Verursachen von E.

Die kausale Einflussnahme von strukturierenden Ursachen erfolgt laut Dretske über Lernprozesse. Denn bei der Art und Weise wie Verhaltensmuster erlernt werden spielen inhaltliche Komponenten eine entscheidende Rolle. Der ausschlaggebende Faktor, der darüber entscheidet, ob etwas als Ursache rekrutiert wird oder nicht ist dessen repräsentationaler Gehalt. Dieser dient als Indikator, ob sich etwas als Ursache eignet oder nicht.

Alles in allem löst Dretskes Dual-Explanandum Theorie die in dieser Arbeit dargestellten Probleme mentaler Kausalität auf eine sehr elegante Art und Weise: so erlaubt diese Theorie mentalen Eigenschaften einerseits kausal relevant zu sein, andererseits verhindert sie aber deren kausale Exklusion durch physikalische Eigenschaften, sowie die mit dem Externalismus einhergehenden Schwierigkeiten für das Verständnis mentaler Kausalität. Während die intrin-

sischen, i.e. neuro(bio)logischen, Eigenschaften eines bestimmten Hirnzustandes die Körperbewegungen auslösen, sind dessen mentale Eigenschaften maßgeblich daran beteiligt es zu der Art Verhalten zu machen, die es ist.

Das Interaktionsproblem

Philosophiegeschichtlich stellt sich das Problem einer adäquaten Erklärung mentaler Verursachung zuerst als sogenanntes Interaktionsproblem dar. Der locus classicus für dieses erste Auftauchen eines Problems mentaler Verursachung in der Philosophie der Moderne ist Descartes' Theorie über die Natur des Menschen: seinem Ansatz zufolge besteht der Mensch gleichermaßen aus zwei Substanzen³³, die zwar grundverschiedenen sind, die aber dennoch ein einheitliches Ganzes bilden und die schließlich auch in kausale Wechselwirkung miteinander treten können. Dementsprechend wird die von Descartes vertretene Position gemeinhin auch als *interaktionistischer (Substanz-)Dualismus*³⁴ bezeichnet.

Descartes' Theorie von Körper und Geist

Die beiden Substanzen sind der Körper (*res extensa*) einerseits und die Seele oder der Geist (*res cogitans*) andererseits. Descartes schreibt beiden Substanzen unterschiedliche, in gewissem Sinne sogar entgegengesetzte, Eigenschaften zu: wie bereits die lateinischen Bezeichnungen vermuten lassen ist die Haupteigenschaft des Körpers die der Ausgedehntheit, wohingegen das zentrale Wesenmerkmal des Geistes im Denken besteht. Wobei der Begriff ‚Denken‘ bei Descartes sehr weit gefasst ist: so gehören beispielsweise auch (Sinnes-)Wahrnehmungen, genauso wie Gefühle oder Empfindungen und schließlich auch Vorstellungen und Denkprozesse zum kartesischen Begriff des Denkens. In zeitgenössischer Nomenklatur würde man daher wohl eher von ‚Bewusstsein‘ als von ‚Denken‘ sprechen – und zwar sowohl im phänomenalen als auch im intentionalen Sinn von Bewusstsein. In seinen *Prinzipien der Philosophie* definiert Descartes den Begriff selbst wie folgt: „Unter Denken verstehe ich alle Vorgänge in unserem Bewusstsein, solange und sofern sie uns bewußt sind.

³³ ‚Substanz‘ wird hier weder im umgangssprachlichen Sinn als (chemischer) Stoff, noch im medizinischen Sinne als homogene Gewebsstruktur (z.B. substantia nigra), noch als ein eine gewisse Wirkung auslösender Stoff (z.B. toxische S.) verstanden. Vielmehr ist der Begriff im philosophischen Sinne zu verstehen: demnach sind Substanzen selbstständig existierende Objekte, die durch ihre Eigenschaften charakterisiert werden. Allerdings sind Substanzen nicht bloße Bündel von Eigenschaften, sondern sie sind vielmehr die Dinge, die eben diese Eigenschaften besitzen. Vgl. Beckermann, Ansgar: *Analytische Einführung in die Philosophie des Geistes*, 2. überarb. Auflage, Berlin, de Gruyter, 2001, S. 6f.

³⁴ Im Folgenden werden die Probleme des interaktionistischen Dualismus hinsichtlich des Phänomens der mentalen Verursachung paradigmatisch an Descartes' Ansatz herausgearbeitet. Aber auch modernere Theorien, die einen interaktionistischen Dualismus vertreten (wie z.B. Poppers & Eccles' Drei-Welten-Lehre), stehen vor denselben bzw. sehr ähnlichen Problemkomplexen. Dualistische Position, die eine Interaktion negieren, haben natürlich kein Problem mit der Erklärung mentaler Verursachung – denn Kausalrelationen zwischen den unterschiedlichen Substanzen existieren per definitionem nicht (z.B. im sogenannten Parallelismus,...).

Deshalb gehören nicht bloß Verstand, Wille und Phantasie, sondern auch die Wahrnehmung hier zum Denken.³⁵

Die körperliche Substanz ist für Descartes – zumindest theoretisch – unendlich teilbar. Dem Geist kommt nur eine zeitliche, aber keine räumliche Existenz zu und somit ist der Geist auch nicht teilbar.³⁶ Schließlich gehorcht alles Körperliche den Gesetzen der Mechanik. Entsprechend betrachtet Descartes den (menschlichen) Körper gleichsam einer völlig determinierten Maschine: „So kann ich auch den menschlichen Körper als eine Art Maschine ansehen, die aus Knochen, Nerven, Muskeln, Adern, Blut und Haut zusammengepaßt ist und auch geistlos all die Bewegungen ausführt, wie sie jetzt unwillkürlich, also ohne den Geist, ablaufen.“³⁷

Trotz ihrer grundlegenden Verschiedenheit bilden Körper und Geist nach Descartes jedoch ein einheitliches Ganzes, er spricht von einer „(...) Verquickung der Seele mit dem Körper (...)“³⁸. So sei der Geist dem „(...) Leibe nicht nur zugestellt wie ein Schiffer dem Schiff, sondern sei (...) aufs innigste mit ihm vereint, durchdringe ihn gleichsam und bilde mit ihm ein einheitliches Ganzes.“³⁹ Eine eingehende Betrachtung der Sachlage legt jedoch die Vermutung nahe, dass Descartes diese systematische Einheit postulieren muss, um (kausale) Interaktion zwischen Körper und Geist nicht a priori auszuschließen. Denn wären Körper und Geist vollkommen voneinander getrennt, wäre es kaum ersichtlich, auf welche Art und Weise sie sich gegenseitig kausal beeinflussen könnten. Das Wirken des Geistes auf den Körper und das Wirken des Körpers auf den Geist ist aber phänomenologisch ein *factum brutum* der alltäglichen Erfahrung jedes Menschen. Dessen ist sich auch Descartes bewusst, wenn er diese Tatsache hypothetisch in Frage stellt und aufzeigt was dies zur Folge hätte: „Wie könnte sonst Ich, ein lediglich denkendes Ding, bei einer Verletzung des Körpers Schmerz empfinden? Ich würde jene Verletzung rein geistig wahrnehmen, wie das Auge des Schiffers es wahrnimmt,

³⁵ Zitiert nach Lauth, Bernhard: *Descartes im Rückspiegel: Der Leib-Seele-Dualismus und das naturwissenschaftliche Weltbild*, Paderborn, Mentis, 2006, S. 74.

³⁶ Manche Autoren kritisieren, dass die Beschreibung der mentalen Substanz vornehmlich ex negativo erfolgt (die mentale Substanz ist primär genau dadurch gekennzeichnet, dass sie eben *nicht* materiell, *nicht* räumlich, *nicht* teilbar, etc. ist) und daher im Grunde unverständlich ist. Aber selbst wenn man die Intelligibilität der mentalen Substanz nicht anzweifelt, drängen sich unbequeme Fragen wie diese auf: Gibt es, analog zur Wissenschaft der Materie, auch eine Wissenschaft, die sich mit dem immateriellen, geistigen Stoff beschäftigt und versucht Gesetzmäßigkeiten innerhalb dieser Substanz zu entdecken? Welche Begriffe würde eine solche Wissenschaft verwenden? Usw. Vgl. McGinn, Colin: *The Character of Mind*, Oxford, Oxford University Press, 1982, S. 25.

³⁷ Descartes, René: *Meditationes de Prima Philosophia / Meditationen über die Erste Philosophie*, Lat. Dt. (Gerhart Schmidt), Stuttgart, Philipp Reclam jun., 1986, S. 201/03.

³⁸ Descartes, René: *Meditationes de Prima Philosophia / Meditationen über die Erste Philosophie*, Lat. Dt. (Gerhart Schmidt), Stuttgart, Philipp Reclam jun., 1986, S. 195.

³⁹ Descartes, René: *Meditationes de Prima Philosophia / Meditationen über die Erste Philosophie*, Lat. Dt. (Gerhart Schmidt), Stuttgart, Philipp Reclam jun., 1986, S. 195.

wenn am Schiff etwas zerbricht; und wenn mein Körper Speise oder Trank braucht, so würde ich dies ausdrücklich erkennen und hätte nicht das verworrene Hunger- oder Durstgefühl.“⁴⁰

Facetten des Interaktionsproblems

An diesem Punkt offenbaren sich für Vertreter des Substanzdualismus Probleme das Wesen mentaler Verursachung gemäß ihrer Theorie zu erklären. Denn es stellt sich natürlich sofort die Frage, *ob* es für so radikal unterschiedliche Substanzen wie Descartes sie beschreibt möglich ist tatsächlich in kausale Wechselwirkung miteinander zu treten. Darüberhinaus scheint es als müsste, selbst wenn dies der Fall wäre, weiterhin erklärt werden, *wie* diese Interaktion gedacht werden kann und *warum* sie überhaupt stattfindet. Die Schwierigkeiten, in die der interaktionistische Substanzdualismus bei der Erklärung der kausalen Wechselwirkung zwischen Körper und Geist gerät, resultieren genauer gesagt aus einigen Eigenschaften, die zahlreiche Theoretiker der Kausalität generell zuschreiben. Diese sind beispielsweise die räumliche Lokalisierbarkeit von Ursache und Wirkung, die These von der Notwendigkeit eines kausalen Nexus, die Impuls- und Energieerhaltung im Rahmen von Kausalbeziehungen, sowie im weiteren Sinne die kausale Vollständigkeit der Physik. Jens Harbecke identifiziert im Wesentlichen zwei Probleme, die dem Substanzdualismus à Descartes eine Erklärung psychophysischer Interaktion erheblich erschweren:

(...) Descartes made his interactionist dualism vulnerable to at least two straightforward objections. Firstly, it seems quite unintelligible how mental states which, according to Descartes, have no spatial extension can engage in causal processes that take place in a determined spatio-temporal region. In other words, if causation is essentially a local matter, it is unclear how essentially unlocalized mental states can be causes. Secondly, as it was known already in the days of Leibniz, the direction of movement of matter is covered by physical law as well. Of course, it still remains a logical possibility that the laws of physics covering vectorial conservation have exceptions or do not apply to all physical systems. However, the empirical evidence that we have today overwhelmingly speaks against this assumption. A violation of the completeness principle (...) has come to be considered an untenable claim.⁴¹

Die beiden Aspekte von Kausalrelationen, die ein adäquates Verständnis mentaler Verursachung unter den Vorzeichen des kartesischen Substanzdualismus nach Harbecke erschweren sind einerseits das Postulat der räumlichen Lokalisierbarkeit von Ursache und Wirkung, sowie andererseits die Forderung nach Impulserhaltung im Rahmen von Kausalbeziehungen. Das Prinzip der Impulserhaltung fußt wiederum maßgeblich auf dem Dogma der Vollständigkeit der Physik.

⁴⁰ Descartes, René: *Meditationes de Prima Philosophia / Meditationen über die Erste Philosophie*, Lat. Dt. (Gerhart Schmidt), Stuttgart, Philipp Reclam jun., 1986, S. 195.

⁴¹ Harbecke, Jens: *Mental Causation – Investigating the Mind's Powers in a Natural World*, Frankfurt, Ontos Verlag, 2008, S. 106

Räumliche Lokalisierbarkeit und kausaler Nexus

Wenn die räumliche Lokalisierbarkeit von Ursache und Wirkung eine Prämisse für jedwede Kausalrelation ist, dann ist völlig unklar wie mentale Ereignisse oder Zustände, die Descartes' Definition zufolge keine räumliche Ausdehnung besitzen, an kausalen Prozessen (die in einem genau eingrenzbaaren Abschnitt der Raumzeit stattfinden) beteiligt sein können. Versteht man Kausalität nämlich im Wesentlichen als eine *lokale* Angelegenheit, dann erscheint es unmöglich, dass mentale Zustände, deren zentrales Charakteristikum gerade ihre Nicht-Räumlichkeit ist, Ursachen oder Wirkungen sind. Somit wäre das Mentale, wie Descartes es beschreibt, *per se* aus der Menge der möglichen Ursachen und Wirkungen ausgeschlossen und ergo wäre auch die Annahme der Existenz mentaler Verursachung von vornherein absurd.

Eine andere Formulierung desselben (oder zumindest eines sehr ähnlichen) Gedankens ist die Behauptung der Notwendigkeit eines kausalen Nexus. Diese Forderung liegt implizit bereits zeitgenössischer Kritik an Descartes' Theorie zugrunde.⁴² Getreu dieser Auffassung ist Kausalität im Kern ihres Wesens ‚mechanistisch‘, d.h. Ursachen ‚treiben‘ ihre Wirkungen stets in gewisser Weise ‚an‘. Zu diesem Zweck ist entweder Berührung oder Ausdehnung – also räumlicher Kontakt – von Ursache und Wirkung vonnöten. Da die geistige Substanz aber per definitionem nicht räumlich ausgedehnt ist, bliebe mentale Verursachung unter diesen Umständen obskur. Dieser Einwand bringt eine populäre Annahme über Kausalität zum Ausdruck, deren Essenz in folgender These ihren Niederschlag findet:

(KN) Jede Kausalrelation benötigt einen Nexus – also eine Schnittstelle oder Verbindung mittels derer die Kausalrelation zu Stande kommt.⁴³

Zumeist wird der für die Verursachung notwendige Nexus im ‚körperlichen‘ (i.e. räumlichen) Kontakt zwischen Ursache und Wirkung gesehen. Aber selbst für den Fall, dass die kausale Schnittstelle nicht im körperlichen Kontakt liegt, müsste ein interaktionistischer Dualist erklären wo sie stattdessen zu suchen ist.⁴⁴ Colin McGinn sieht darin ein nahezu aussichtsloses Unterfangen:

⁴² Vgl. z.B. erstes Zitat aus dem Briefwechsel zwischen Prinzessin Elisabeth und Descartes unten.

⁴³ Vgl. Heil, J. & Robb, D.: ‚Mental Causation‘ in: SEP

⁴⁴ Einige Textstellen deuten darauf hin, dass Descartes ein solches mechanistisches Verständnis von Kausalität teilt. Nach seiner Theorie des Geistes findet psychophysische Wechselwirkung nur mit wesentlicher Beteiligung des Gehirns statt.

(...) on a dualist view, there is the notorious difficulty of accounting for causal interaction between mental and physical events if they are to be located in such diverse substances. We generally conceive of causal interaction as proceeding via some sort of mechanism, in such a way that the interacting things engage with each other in some intelligible nexus. But this sort of intelligible connection is precisely what is lacking on the dualist account of mind-body interaction, since the very point of that account is to insist on the radical difference of nature between mental and physical phenomena. Try to imagine what sort of mechanism might enable material and immaterial substance to come causally together: in so far as you have any conception of the nature of immaterial substance, this must seem a hard task – certainly we cannot legitimately appeal to the sorts of causality mediated by the physical forces studied in the sciences of matter.⁴⁵

Üblicherweise gehe man also davon aus, dass kausale Interaktion funktioniere, prinzipiell getreu eines gewissen Mechanismus', der garantiere, dass Ursache und Wirkung auf nachvollziehbare Weise an einer Schnittstelle ineinandergreifen, so McGinn. Allerdings sei eine solche plausible Verbindung im Falle der kausalen Wechselwirkung zwischen Körper und Geist im Substanzdualismus Descartescher Prägung nicht vorhanden. Schließlich postuliere diese Theorie des Geistes gerade eine grundlegende Verschiedenartigkeit von Mentalem und Physikalischem. Welche Art von Mechanismus es in diesem Rahmen erlauben würde, materielle und immaterielle Substanzen kausal zusammenzuführen, sei allerdings völlig unklar. Jedenfalls könne diese Erklärungslücke nicht von der Form von Kausalität, die aus der Physik bekannt ist, geschlossen werden. Wie genau allerdings mentale Kausalität ausgestaltet sein sollte, wenn sie völlig divergent zu physikalischer Kausalität ist, darüber schweigen sich interaktionistische Substanzdualisten jedoch zumeist aus, so McGinn.

Impuls- bzw. Energieerhaltung und Vollständigkeit der Physik

Einen weiteren Problemkomplex für das Verständnis mentaler Verursachung im Substanzdualismus stellen drei fundamentale Axiome der Naturwissenschaften dar – und zwar Impuls- und Energieerhaltung, sowie Vollständigkeit⁴⁶ der Physik. Dem ersten Diktum zufolge bleibt der Gesamtimpuls in einem abgeschlossenen System stets konstant. Dies spielt insbesondere in der Beschreibung von Stoßprozessen eine gravierende Rolle und hat zur Folge, dass der Gesamtimpuls aller am Stoß beteiligten Objekte vor und nach dem Aufprall genau gleich groß ist. Trifft beispielsweise eine Billardkugel eine andere, ruhende Kugel direkt und überträgt so ihren Gesamtimpuls vollständig auf jene, dann bleibt die erste Kugel liegen und die zweite rollt mit derselben Geschwindigkeit weiter, die die erste Kugel vor dem Aufprall hatte (wenn man von den Verlusten durch Reibung einmal absieht). In Hinblick auf mentale Kausalität im

⁴⁵ McGinn, Colin: *The Character of Mind*, Oxford, Oxford University Press, 1982, S. 27.

⁴⁶ In der Literatur taucht dieses Prinzip oder besser gesagt die ihm zugrundeliegende Idee unter mehreren Namen auf. So ist beispielsweise gelegentlich auch von der ‚kausalen Abgeschlossenheit‘ oder der ‚explanatorischen Autonomie oder Autarkie‘ der Physik die Rede.

Substanzdualismus ist die Impulserhaltung insofern problematisch, als dass sie natürlich auch für Körperbewegungen gelten muss. Wenn die besagten Körperbewegungen allerdings eine mentale Ursache besitzen, dann müsste ein kartesischer Substanzdualist erklären wie der entsprechende Impuls von der mentalen Ursache auf die physikalische Wirkung übertragen werden kann. Im Falle rein physikalischer Verursachung wie etwa bei der oben beschriebenen Kollision zweier Billardkugeln ist es offensichtlich wie und wo die erste Kugel ihren Impuls auf die zweite überträgt. Bei mentalen Ursachen von Körperbewegungen hingegen ist es alles andere als klar wie und wo eine Impulsübertragung stattfinden könnte, schließlich besitzt die mentale Ursache keine räumliche Ausdehnung. Daher erscheint es auch wenig sinnvoll nach einem konkreten Ort Ausschau zu halten, an dem die Impulsübertragung erfolgen könnte. Genauso absurd wäre es jedoch anzunehmen, der sich bewegende Körper erführe eine Veränderung seines Impulses sozusagen aus dem Nichts.

Ein analoges Argument lässt sich auch aus dem Energieerhaltungssatz ableiten. Gemäß des Energieerhaltungssatzes ist nämlich die Gesamtmenge der Energie in einem geschlossenen System notwendigerweise konstant. Obwohl die Energie in einem Systems zwar durchaus in unterschiedliche Formen überführt werden kann (z.B. von Bewegungsenergie in Wärmeenergie), kann *innerhalb* dieses Systems weder Energie erzeugt noch vernichtet werden. Da der menschliche Körper offensichtlich einen Teil des physikalischen Gesamtsystems repräsentiert, scheint somit die Möglichkeit ausgeräumt, dass etwas, das nicht Teil der physikalischen Ordnung ist – wie z.B. mentale Zustände nach Descartes – als (alleinige) Ursache etwa für Körperbewegungen in Frage kommt. Schließlich ist anzunehmen, dass bei allen Kausalitätsprozessen entweder ein Energieübertrag von der Ursache zur Wirkung stattfindet oder dass die Ursache zumindest eine Umwandlung der Energieform bei der Wirkung zur Folge hat.

Die Annahme der Vollständigkeit der Physik ist wiederum eine Prämisse für den Energieerhaltungssatz – schließlich bezieht sich die Kernaussage des Energieerhaltungssatzes auf *geschlossene* physikalische Systeme. Im Wesentlichen besagt die These von der Vollständigkeit der Physik, dass es für jede physikalische Wirkung eine hinreichende physikalische Ursache gibt. Demzufolge ist es offensichtlich prinzipiell unmöglich, dass mentale Ursachen physikalische Wirkungen haben – es sei denn man zieht eine systematische Überdeterminierung solcher Wirkungen als eine vernünftige Option in Betracht. So könnte beispielsweise eine Körperbewegung daher niemals durch einen Wunsch verursacht worden sein. Vielmehr müsste die besagte Körperbewegung eine rein physikalische Ursache besitzen, deren Auftreten *alleine* schon hinreichend für den Ortswechsel des entsprechenden Körper(teil)s gewesen wäre. In

anderen Worten könnte eine mentale Ursache keinen essentiellen Beitrag zur Verursachung der Körperbewegung leisten – schließlich würde diese selbst dann erfolgen, wenn die mentale Ursache nicht eingetreten wäre.

Descartes' Erklärung psychophysischer Interaktion innerhalb seiner Theorie

Descartes geht davon aus, dass die eigentliche Interaktion zwischen Körper und Geist im oder zumindest unter federführender Mitwirkung des Gehirns stattfindet. So stellt sich Descartes die kausale Geschichte einer Schmerzempfindung in etwa wie folgt dar: „Werden z.B. die Nerven des Fußes heftig und in ungewohnter Weise erregt, so pflanzt sich diese Erregung durch das Rückenmark bis ins Innere des Gehirns fort und gibt dort dem Geist das Zeichen zu einer Empfindung, nämlich zu der im Fuß lokalisierten Schmerzempfindung. Dadurch wird nun der Geist angetrieben, die Ursache des Schmerzes als etwas dem Fuß Schädliches nach Kräften zu beseitigen.“⁴⁷ Descartes ist sich also bereits durchaus bewusst, dass das Gehirn in Sachen psychophysischer Interaktion eine Schlüsselrolle inne hat, denn Schmerzen im Fuß können beispielsweise auch auftreten, ohne dass die Quelle des Schmerzes tatsächlich im Fuß zu lokalisieren wäre: „Wenn nämlich eine Ursache nicht im Fuß, sondern an einer andern Stelle der Nerven auf ihrem Weg vom Fuß zum Gehirn, oder auch im Gehirn selbst, genau die nämliche Erregung hervorrufen würde, die bei einer Verletzung des Fußes einzutreten pflegt, so werden wir doch den Schmerz gleichsam im Fuß empfinden, und unsere Empfindung irrt dann auf natürliche Weise. Denn ein und dieselbe Gehirnerregung kann dem Geist stets nur die gleiche Empfindung mitteilen (...)“⁴⁸

Descartes geht offensichtlich davon aus, dass es für die kausale Interaktion zwischen Körper und Geist sozusagen einer Kommunikation zwischen beiden Substanzen bedarf. Dieses ‚Kommunikationsinterface‘ ist nun im Gehirn, genauer gesagt in der Epiphyse (Zirbeldrüse), zu verorten. Nach Descartes' Vorstellung gleicht das Nervensystem einem feingliedrigen Röhrensystem im Körper des Menschen. Durch die Nervenröhren bewegen sich die so genannten Lebensgeister (*spirti animali*), die Reiz-Informationen durch den Körper ins Gehirn und umgekehrt vom Gehirn in die Peripherie transportieren. In der Zirbeldrüse kommt es zur kausalen Wechselwirkung zwischen Körper und Geist: bestimmte Gehirnerregungen teilen

⁴⁷ Descartes, René: *Meditationes de Prima Philosophia / Meditationen über die Erste Philosophie*, Lat. Dt. (Gerhart Schmidt), Stuttgart, Philipp Reclam jun., 1986, S. 209.

⁴⁸ Descartes, René: *Meditationes de Prima Philosophia / Meditationen über die Erste Philosophie*, Lat. Dt. (Gerhart Schmidt), Stuttgart, Philipp Reclam jun., 1986, S. 211.

dem Geist gewisse Empfindungen mit, diese wiederum rufen bestimmte Bewegungen des Körpers hervor.

Der Einwand, dass das Auftreten psychophysischer Wechselwirkung aufgrund der fundamentalen Verschiedenheit der beiden Substanzen mysteriös bleibe, wurde bereits von einigen Zeitgenossen Descartes' geäußert (vgl. Lauth, Kapitel 9). Schließlich gehören beide Substanzen per definitionem unterschiedlichen und strikt voneinander getrennten Seinsbereichen an. Aus diesem Grund ist es unklar, ob überhaupt kausale Wechselwirkung zwischen beiden Domänen stattfinden kann und wie man sich eine solche Interaktion vorstellen könnte. Gassendi etwa schreibt: „Wie kann die Seele den Körper bewegen, wenn sie selbst gar nicht materiell ist (...)“⁴⁹ Allerdings geht Descartes zunächst nicht auf die geäußerte Kritik ein. Erst im Briefwechsel mit Prinzessin Elisabeth von Böhmen bezieht Descartes Stellung. Prinzessin Elisabeth schreibt im Mai 1643 in einem Brief an Descartes:

Wie kann die Seele des Menschen die Lebensgeister dazu veranlassen, die Willkürhandlungen auszuführen (da sie doch nur eine denkende Substanz ist)? Denn es scheint, daß jede Bewegung durch einen Stoß verursacht wird, wobei die Art des Stoßes von den Eigenschaften und der Form der Oberfläche abhängt, durch den der Stoß ausgeführt wird. In den ersten beiden Fällen wird Berührung vorausgesetzt und beim dritten räumliche Ausdehnung. Sie schließen aber diese vollständig aus dem Begriff aus, den Sie von der Seele haben, und jene erscheint mir unvereinbar mit einem immateriellen Gegenstand. Deshalb bitte ich Sie um eine spezifischere Definition der Seele als in ihrer Metaphysik (...).⁵⁰

In seinem Antwortschreiben geht Descartes schließlich auf die von Prinzessin Elisabeth geäußerte Kritik ein. Er gesteht ein, dass er Körper und Seele bislang hauptsächlich jeweils anhand einer ‚notion primitive‘ kontrastiert habe, wobei eine ‚notion primitive‘ einen nicht weiter definier- und erklärbaren Grundbegriff des Denkens darstellt. Der besagte Grundbegriff für den Körper ist die Ausgedehntheit und der für die Seele ist das Denken. Dabei sei es ihm sehr daran gelegen gewesen zu zeigen, dass die beiden Haupteigenschaften, der von ihm postulierten Substanzen, grundverschieden und unvereinbar sind. Ein weiterer fundamentaler Schlüsselbegriff jedoch, der sowohl den Körper als auch die Seele betrifft, wurde Descartes zufolge in seiner Beschreibung bislang vernachlässigt – und zwar der ihrer Vereinigung (union). Dieser Begriff der Verbindung von Körper und Seele ist schließlich essentiell für das Verständnis psychophysischer Wechselwirkung. Entsprechend beantwortet Descartes Elisabeths Brief:

Erstens erwäge ich, daß wir gewisse primitive Begriffe haben, die gleichsam ursprünglich sind, nach deren Modell wir alle unsere anderen Erkenntnisse bilden. Und es gibt nur sehr wenig derartige Begriffe;

⁴⁹ Zitiert nach Lauth, Bernhard: *Descartes im Rückspiegel: Der Leib-Seele-Dualismus und das naturwissenschaftliche Weltbild*, Paderborn, Mentis, 2006, S. 187

⁵⁰ Zitiert nach Lauth, Bernhard: *Descartes im Rückspiegel: Der Leib-Seele-Dualismus und das naturwissenschaftliche Weltbild*, Paderborn, Mentis, 2006, Kapitel 9

denn nach den allgemeinsten des Seins, der Zahl, der Dauer, usw., die allem zukommen, was wir begreifen können, haben wir für den Körper im Besonderen nur den Begriff der Ausdehnung, aus dem die Begriffe der Gestalt und der Bewegung folgen; und für die Seele allein haben wir den des Gedankens, in dem die Wahrnehmungen des Begriffsvermögens und die Neigungen des Willens einbegriffen sind; für Seele und für Körper zusammen haben wir schließlich nur den ihrer Vereinigung, von welchem die Kraft abhängt, die die Seele hat, den Körper zu bewegen, und der Körper, auf die Seele zu wirken, indem er deren Gefühle und Leidenschaften verursacht.⁵¹

Diese Erklärung der kausalen Interaktion zwischen Körper und Geist ist allerdings nicht unbedingt sehr befriedigend. Denn zusammenfassend könnte Descartes' Argumentationsgang in etwa folgendermaßen dargestellt werden:

- (1) Die Verbindung von Körper und Seele erklärt die kausale Wechselwirkung zwischen den beiden Substanzen.
- (2) Die Verbindung von Körper und Seele ist eine ‚notion primitive‘, d.h. ein Grundbegriff des Denkens.
- (3) Grundbegriffe des Denkens sind per definitionem nicht weiter definier- und erklärbar.
- (4) Daher lässt sich nicht mehr über die psychophysische Wechselwirkung zwischen Körper und Seele sagen.

So ist es nicht weiter verwunderlich, dass diese Erklärung auch Prinzessin Elisabeth nicht zufrieden stellen konnte. In einem späteren Brief an Descartes schreibt sie, sie könne weiterhin nicht verstehen,

(...) wie die Seele (die unausgedehnt und immateriell ist) den Körper bewegen kann... Und ich gestehe, daß es mir leichter fallen würde, der Seele eine Materie und eine Ausdehnung zuzuschreiben, als einem immateriellen Wesen die Fähigkeit, einen Körper zu bewegen und von ihm bewegt zu werden. Denn, wenn das Erstere durch Information geschähe, müßten die Lebensgeister, die die Bewegung verursachen, intelligent sein, was Sie aber keinem körperlichen Wesen zubilligen. Und obwohl Sie in ihren metaphysischen Meditationen die Möglichkeit des Zweiten nachweisen, bleibt es gleichwohl sehr schwer zu verstehen, wie eine Seele, so wie Sie sie beschrieben haben, nachdem sie die Fähigkeit und die Gewohnheit des folgerichtigen Denkens erworben hat, dies alles durch einige Dünste verlieren kann und wie sie, die doch ohne Körper existieren kann und mit ihm nichts gemeinsam hat, dermaßen von ihm beherrscht werden kann.⁵²

In seiner Erwiderung auf Prinzessin Elisabeths Brief hält Descartes an seiner Theorie der undefinierbaren Grundbegriffe fest. Um das Thema endgültig zu erledigen, weist er schließlich darauf hin, dass die Verbindung von Körper und Seele in unserem Alltagserleben völlig unproblematisch ist. Schlussendlich erfahren wir jeden Tag aufs Neue wie der Körper kausal auf die Seele wirkt oder wie die Seele etwas im Körper verursacht. „Daher kommt es, daß dieje-

⁵¹ Zitiert nach Lauth, Bernhard: *Descartes im Rückspiegel: Der Leib-Seele-Dualismus und das naturwissenschaftliche Weltbild*, Paderborn, Mentis, 2006, S. 188.

⁵² Zitiert nach Lauth, Bernhard: *Descartes im Rückspiegel: Der Leib-Seele-Dualismus und das naturwissenschaftliche Weltbild*, Paderborn, Mentis, 2006, S. 189.

nigen, die niemals philosophieren und sich nur ihrer Sinne bedienen, nicht daran zweifeln, daß die Seele den Körper bewegt, und daß der Körper auf die Seele wirkt; sie betrachten aber beide als eine einzige Sache, das heißt, sie begreifen nur ihre Vereinigung; denn die zwischen zwei Dingen bestehende Vereinigung begreifen heißt, sie als ein einziges begreifen.“⁵³

Aber selbst wenn man diese Beweisführung akzeptiert, so erklärt sie nur, *dass* Seele und Körper miteinander verbunden sind, aber die eigentlich interessanteren Fragen nach dem *Wie* und *Warum* der kausalen Interaktion bleiben davon gänzlich unberührt. Dementsprechend fasst Prinzessin Elisabeth die Ergebnisse ihres Briefwechsels mit Descartes wie folgt zusammen:

Ich finde auch, daß die Sinne mir zeigen, daß die Seele den Körper bewegt, daß sie mir aber keineswegs die Art und Weise zeigen (ebenso wenig wie der Verstand oder das anschauliche Vorstellungsvermögen), wie sie dies tut. Und aus diesem Grund denke ich, daß es in der Seele Eigenschaften gibt, die uns unbekannt sind, und die vielleicht alles über die Unausgedehntheit der Seele umstürzen könnten, wovon mich Ihre Meditationen mit so guten Argumenten überzeugt haben.⁵⁴

Weitere Lösungen des Interaktionsproblems aus Sicht des Substanzdualismus

Lösungen für das Problem der räumlichen Lokalisierbarkeit und das Nexus-Problem

Um die Schwierigkeiten zu vermeiden, die dabei entstehen, wenn man den Nexus für die Kausalbeziehung zwischen Körper und Geist erklären muss, könnte ein interaktionistischer Substanzdualist natürlich auch schlichtweg die These (KN) zurückzuweisen und eine bloße Regularitätstheorie von Kausalität vertreten, derzufolge Ursache und Wirkung eines Ereignisses keiner Schnittstelle bedürfen, an der oder mittels derer die Interaktion stattfindet; vielmehr wäre eine Kausalitätsbeziehung vielmehr als eine konstante Konjunktion zweier Ereignistypen zu verstehen.

Eine Argumentationsstrategie, die in etwa diese Richtung einschlägt, behauptet es gebe andere Formen von Kausalitätsbeziehungen, wie z.B. die der Gravitation, bei denen auch unklar sei, ob sie eines Kontaktes zwischen Ursache und Wirkung bedürfen. Denn zwischen einem schweren Körper und einem von ihm angezogenen leichteren Körper besteht offensichtlich kein räumlicher oder körperlicher Kontakt; nichtsdestotrotz findet zwischen beiden Körpern offensichtlich kausale Wechselwirkung statt. Folglich kann man die kausale Interaktion zwi-

⁵³ Zitiert nach Lauth, Bernhard: *Descartes im Rückspiegel: Der Leib-Seele-Dualismus und das naturwissenschaftliche Weltbild*, Paderborn, Mentis, 2006, S. 189f.

⁵⁴ Zitiert nach Lauth, Bernhard: *Descartes im Rückspiegel: Der Leib-Seele-Dualismus und das naturwissenschaftliche Weltbild*, Paderborn, Mentis, 2006, S. 190.

schen Körper und Geist nicht allein deshalb zurückweisen, weil sie ohne Kontakt stattfindet oder weil unklar ist wie sie funktioniert. Dazu äußert sich Dardis wie folgt:

Obviously there is no question that gravity operates on things. To this day, though, gravity continues to be a mystery, and still we don't know whether it operates by contact. So the facts that the mind-body causation apparently doesn't involve contact that we don't understand how it works cannot constitute an objection to mind-body causation.

It is easy, in a sense, to explain the workings of things that have a complex and easily accessible structure. You just have to know the detailed structure of the things involved and their particular causal contributions. How does the sunlight cause plants to flourish? Electromagnetic radiation of a certain wavelength causes changes in chlorophyll molecules. How does that causation work? Photons of a certain energy interact with certain portions of the molecules. How does that causation work? Well, that depends on whether those interactions still have structure. If they do, then we continue explaining in terms of ever-finer structure. But there comes a point at which the structure is exhausted. The things that interact have no more structure. The causation is brute. There is nothing more to say than "this is what happens".

Brute facts are the facts in terms of which other facts are explained (...). They themselves cannot be explained. Explanation must stop somewhere, and hence some facts must remain brute.⁵⁵

Dementsprechend schlägt Dardis auch vor, Descartes solle auf nachstehende Weise auf Prinzessin Elisabeths Einwand reagieren: „So Descartes can answer Elisabeth's question in the following way: there is mind-body interaction. There is no explanation for it. But that is not a difficulty for Cartesian interaction. Every account of the world must have a place for kinds of causation for which there is no explanation.“⁵⁶ Diese Strategie basiert auf dem Gedanken, dass Kausalität manchmal sozusagen ‚nackt‘ oder ‚direkt‘ ist, d.h. nicht in jedem Fall einer Verbindung zwischen Ursache und Wirkung bedarf, die die Verursachung erklärt. Auch dieser Argumentationsgang kommt somit dem Zurückweisen der oben beschriebenen These (KN) gleich: es gibt zumindest einige Fälle, in denen Ursache und Wirkung ohne einen vermittelnden Nexus zustande kommen. Sicherlich ist dies eine durchaus legitime Argumentationsweise für Substanzdualisten: Neben der mechanistischen Kausalität, die im materiellen Gegenstandsbereich kausal wirksam ist, existiert auch mentale Kausalität und diese erklärt die kausale Wechselwirkung zwischen Körper und Geist. Mentale Kausalität selbst ist in diesem Fall jedoch eine so fundamentale Beziehung in der Welt, dass sie nicht weiter erklärt werden kann. Somit wäre die Frage wie Geist und Körper gegenseitig aufeinander wirken können schlichtweg sinnlos. Oder wie Richardson bemerkt: „So long as Descartes holds (...) that there are two fundamental and irreducible forms of causal interaction, there can be no sense given to the question of how mind acts on body.“⁵⁷

⁵⁵ Dardis, Anthony: *Mental Causation – The Mind-Body Problem*, New York, Columbia University Press, 2008, S. 34.

⁵⁶ Dardis, Anthony: *Mental Causation – The Mind-Body Problem*, New York, Columbia University Press, 2008, S. 34.

⁵⁷ Richardson, R. C.: „The ‚Scandal‘ of Cartesian Interactionism“ in: *Mind*, Volume 91, 1982, S. 26.

Zusammenfassend lässt sich sagen, dass Descartes' Argumentation hinsichtlich der Möglichkeit einer kausalen Wechselwirkung zwischen Körper und Seele nicht wirklich stichhaltig ist. Sein erstes Argument ist, dass die Vereinigung von Körper und Seele die kausale Interaktion erklärt. Allerdings deklariert er die besagte Vereinigung als einen nicht weiter definierbaren Grundbegriff des Denkens. Somit endet die Argumentation an dieser Stelle. Es ist offensichtlich, dass dieses Argument den faden Beigeschmack hat eine bloße Ad hoc-Lösung für ein unangenehmes Problem zu sein. Auch das zweite Argument Descartes' ist nicht sehr schlagkräftig: demzufolge bedarf die kausale Wechselwirkung zwischen Körper und Seele keiner weiteren Erklärung, weil sie eine offensichtliche Tatsache unserer alltäglichen Lebenswelt ist. Gerade im Hinblick auf den radikalen Skeptizismus hinsichtlich der Alltagserfahrung, den Descartes an anderer Stelle vertritt (vgl. sein Dämon- und sein Traumargument in den *Meditationes*), erscheint auch dieses Argument in einem fadenscheinigen Licht. Insgesamt betrachtet, besitzen die beiden einzigen Argumente, die Descartes für das Vorhandensein psychophysischer Wechselwirkung vorbringt, also kaum Überzeugungskraft.

Schließlich können Descartes' Ausführungen, selbst wenn man seiner Argumentation folgt, nur erklären, *dass* es kausale Interaktion zwischen Körper und Seele gibt. Gesteht man das zu, schließen sich aber unweigerlich die Fragen an *wie* man sich diese kausale Wechselwirkung vorstellen muss und *weshalb* sie überhaupt stattfindet. Zu diesen Problemen schweigt sich Descartes allerdings gänzlich aus.

Von einem logischen Standpunkt aus, ist seine Beweisführung einwandfrei, aber natürlich schließt sich sofort die Frage an, inwiefern sie überhaupt irgendetwas erklären kann. So sehen manche den Niedergang des interaktionistischen Dualismus eben darin begründet, dass er auf diese Frage keine Antwort geben kann – so glaubt Kim: „(...) that its inability to account for mental causation was the downfall of Descartes's mind-body dualism.“⁵⁸ Und selbst für den Fall, dass man diesen Gedankengang akzeptiert, bleiben weitere gravierende Probleme hinsichtlich mentaler Verursachung für den interaktionistischen Dualismus bestehen.

⁵⁸ Kim, Jaegwon: *Physicalism, or something near enough*, Princeton University Press, Princeton, 2005, S. 73.

Das Pairing Problem

Das sogenannte Pairing Problem⁵⁹ stellt den interaktionistischen Dualismus vor die Frage, wie man sich genau die gegenseitige Zuordnung von Mentalem und Physikalischem vorzustellen habe. Welche Funktion in diesem Zusammenhang eine eindeutige Verknüpfung beider Bereiche erlaube ist unklar, so Jaegwon Kim. Die Essenz des Pairing Problems beschreibt er mit folgendem Gedankenexperiment:

There are two souls, A and B, and they perform an identical mental act at time *t*, as a result of which a change occurs in material substance M shortly after *t*. We may suppose that mental actions of the kind involved generally cause physical changes or the sort that happened in M, and, moreover, that in the present case it is soul A's action, not soul B's, that caused the change in M. Surely, such a possibility must exist. But ask: What relation might serve to pair soul A's action with the change in M, a relation absent in the case of soul B's action and the change in M? That is, what could be the pairing relation in this case? Evidently, no spatial relations can be invoked to answer the question, for souls are not in space and are not able to bear spatial relations to material things. Soul A cannot be any "nearer" to material object M, or more propitiously "oriented" in relation to it, than soul B is. Is there anything that can do for souls what space, or a network of spatial relations, does for a material thing?⁶⁰

Um dieses Gedankenexperiment ein wenig mit Leben zu füllen, ein konkretes Beispiel: Nehmen wir mit Descartes an, es gäbe immaterielle Seelen, die kausal mit ‚ihren‘⁶¹ jeweiligen Körpern in Wechselwirkung treten können. Stellen wir uns nun weiter vor, dass zwei Seelen und zwar Franz' Seele und Fritz' Seele gleichzeitig einen identischen Wunsch verspüren – nämlich den rechten Arm ihres (jeweiligen) Körpers anzuheben. Und tatsächlich heben unmittelbar nachdem sich der Wunsch in Franz' und Fritz' Seele herausgebildet hatte, auch die Körper von Franz und Fritz gleichzeitig ihren rechten Arm an. Das Pairing Problem weist nun auf den Umstand hin, dass im interaktionistischen Dualismus nicht erklärt werden kann warum gerade der Wunsch von Franz' Seele kausal dafür verantwortlich sein soll, dass Franz' Körper den rechten Arm anhebt und nicht der Wunsch von Fritz' Seele (und umgekehrt). Welche Relation verbindet also Franz' Seele kausal mit Franz' Körper und Fritz' Seele kausal mit Fritz' Körper? Offensichtlich kann es keine räumliche Beziehung sein – also Franz' Seele kann nicht einfach ‚näher‘ an Franz' Körper sein oder etwa günstiger auf ihn oder zu ihm ausgerichtet sein, denn schließlich ist Franz' Seele per definitionem nicht räumlich und somit ist es auch sinnlos zu sagen, sie sei näher an (oder besser ausgerichtet auf/zu) Franz' Körper

⁵⁹ Im Folgenden ist immer vom ‚Pairing Problem‘ die Rede – ich bevorzuge die englischsprachige Nomenklatur, weil diese schlichtweg eleganter klingt als eine sperrige Umschreibung im Deutschen oder eine direkte Übersetzung ins Deutsche.

⁶⁰ Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 79f.

⁶¹ Es ist anzunehmen, dass auch interaktionistische Dualisten aus Kohärenzgründen davon ausgehen, dass eine Seele in der Regel nur mit einem – eben ‚ihrem‘ – Körper in kausale Wechselwirkung tritt und dass dieses Gespann aus genau dieser Seele und genau jenem Körper auch über die Zeit hinweg konstant verbunden ist. Ferner scheint es so, als habe Descartes die Vereinigung von Seele und Körper gerade zu dem Zweck so postuliert, um eine konstante Konjunktion einer bestimmten Seele mit einem gewissen Körper zu etablieren.

als Fritz' Seele und deshalb kausal verantwortlich für Franz' Körperbewegung. Aber welche Beziehung ordnet Franz' Seele Franz' Körper kausal zu? Kraft was ist es also wahr, dass Franz' Seele mit Franz' Körper eine kausale Einheit bildet?⁶² Und warum sollte nicht der Wunsch von Fritz' Seele Franz' Armheben verursacht haben? Descartes hatte zu diesem Zweck einfach die Verbindung von Körper und Seele postuliert – allerdings erklärt er die besagte Vereinigung von Körper und Geist nicht weiter, so dass sie letzten Endes mysteriös bleibt.

Das Pairing Problem hängt also unmittelbar damit zusammen, dass mentale Kausalität im interaktionistischen Dualismus als fundamental und somit nicht näher beschreibbar definiert werden muss (wie oben gezeigt wurde, droht der interaktionistische Dualismus ansonsten an der Frage nach dem Wie der kausalen Interaktion zu scheitern). Wie bereits oben erwähnt bleibt das *Wie* mentaler Verursachung (verstanden als fundamentale Beziehung in der Welt) rätselhaft. Während z.B. bei Ereignissen, die physikalischer Kausalität unterliegen, einfach die entsprechende Kausalkette zurückverfolgt werden kann, ist dies bei mentaler Verursachung im interaktionistischen Dualismus nicht möglich. Ein analoges Beispiel aus dem Bereich der Physik zum obigen ‚Seelenbeispiel‘, könnte man in etwa folgendermaßen skizzieren: aus zwei Gewehren A und B wird gleichzeitig ein Schuss abgegeben und dadurch sterben gleichzeitig zwei Personen. Der Schuss aus Gewehr A tötet Adam, während der Schuss aus der Flinte B Bob tötet.⁶³ In diesem Fall kann das (kausale) Pairing Problem, laut Kim, auf zwei Arten gelöst werden. Erstens kann man eine lückenlose Kausalkette vom Abfeuern eines Geschosses aus A und Adams Tod (entsprechend bei B und Bobs Ableben) verfolgen, während es keine solche fortwährende Kausalkette vom Feuern aus A und Bobs Tod (oder vom Schuss aus B und Adams Tod) gibt. Zweitens wird man, wenn man nach einer Pairing-Beziehung P sucht, die zwischen dem Abfeuern von A und Adams Tod und dem Abfeuern von B und Bobs Tod besteht (und nicht zwischen dem Abfeuern von A und Bobs Tod und umgekehrt) herausfinden, dass es *räumliche* Beziehungen sind, die helfen das Abfeuern von A mit Adams Tod zu verbinden. So war im Moment des Abfeuerns der Waffen z.B. A (und nicht B) in einer angemessenen Entfernung zu Adam und zeigte zugleich in seine Richtung (analog mit B und Bob). Folglich sieht es so aus, als wären es in diesem Fall (und vielleicht in allen Fällen physikalischer Verursachung zwischen unterschiedlichen Gegenständen) *räumliche* Beziehungen, die als Pairing-Beziehung P dienen können. Es ist offensichtlich, dass räumliche Beziehungen

⁶² Diese Frage soll metaphysisch und nicht epistemisch verstanden werden. D.h. hier ist allein von Interesse warum genau *diese* Paarung (also Franz' Seele und Franz' Körper bzw. Fritz' Seele und Fritz' Körper) besteht und nicht ob oder wie wir wissen können, dass genau diese Paarung tatsächlich besteht.

⁶³ Vgl. Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 79.

bei mentaler Kausalität nicht die Rolle einer Pairing-Beziehung P spielen können, denn mentale Ereignisse sind per definitionem nicht räumlich, somit kann die Pairing-Beziehung P zwischen mentalen Entitäten und physikalischen Entitäten nicht räumlicher Natur sein. Dann stellt sich jedoch die Frage wie P sonst beschaffen sein sollte. Eine Frage, auf die der interaktionistische Dualismus Kim zufolge keine bzw. keine befriedigende Antwort geben kann.

Das Pairing Problem ist jedoch nicht, wie zunächst zu erwarten wäre, lediglich ein Problem für mentale Verursachung in Geist-Körper-Richtung, sondern stellt Kim zufolge auch für Geist-zu-Geist-Verursachung ein gravierendes Problem dar:

Consider a purely mental world, a world inhabited only by Cartesian souls; such a world must be possible, since souls are “substances”, that is, independent existents. Soul A acts in a certain way and so does soul B at the same time. This is followed by certain changes in two other souls, A* and B*. Suppose that actions of A and B are causes of the changes in A* and B*. But which cause caused which effect? If we want a solution that is analogous to case (2) above for the firings of guns and the deaths, what we need is a pairing relation *R* such that *R* holds, say, for A and A*, and for B and B*, but not for A and B*, or for B and A*. Since these entities are immaterial souls outside physical space, *R* cannot be, or include, a spatial relation, or any other kind of physical property or relation. The radical nonspatiality of mental substances rules out the possibility of invoking spatial relationships to ground cause-effect pairings.⁶⁴

Die Vorstellung einer rein mentalen, nicht materiellen Welt mag zunächst abwegig klingen, muss aber nach Descartes' Definition der *res cogitans* möglich sein. Schließlich können Körper und Seele unabhängig voneinander existieren und wo sollten Seelen sonst existieren, wenn nicht in einer nicht-stofflichen, mentalen Welt? Zudem wird ein interaktionistischer Dualist in der Tradition von Descartes annehmen, dass Seelen miteinander kausal in Wechselwirkung treten können. Ansonsten wäre beispielsweise jegliche Diskussion (oder gar jedwede Kommunikation) sinnlos – denn, wenn Geist-zu-Geist-Verursachung unmöglich wäre, könnte ein Geist einen anderen nicht durch gute Argumente von einer Sache überzeugen (es wäre auch schlichtweg unmöglich, dass sich zwei Geister gegenseitig wahrnehmen, denn dazu müsste der Eindruck des Einen eine Wahrnehmung im Anderen verursachen).

Kim geht davon aus, dass das Pairing Problem bei Geist-zu-Geist-Verursachung in gleicher Weise auftritt wie bei der Verursachung in Geist-Körper-Richtung. Um beim vorherigen Beispiel zu bleiben: Nehmen wir also an, dass Franz' und Fritz' Seele simultan eine gewisse Handlung vollziehen, etwa ein Argument für den interaktionistischen Dualismus äußern.⁶⁵

⁶⁴ Vgl. Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 80.

⁶⁵ Ich bin mir bewusst, dass es problematisch ist hier von ‚äußern‘ zu sprechen, weil Seelen schließlich keine Mäuler haben mit denen sie etwas aussprechen könnten und keine Ohren mit denen sie die Äußerung der anderen Seele wahrnehmen könnten. Um des Arguments willen bitte ich aber über diesen Makel hinwegzusehen.

Nehmen wir weiterhin an, dass dieses Argument die Seelen von Franziska und Frieda von der Wahrheit des interaktionistischen Dualismus überzeugen soll. Genauer gesagt ist man vielleicht geneigt anzunehmen, dass die Äußerung des Arguments von Franz' Seele Franziskas Seele überzeugt hat, wohingegen es Fritz' Seele war, die Friedas Seele überzeugte. Analog zum Fall der Geist-Körper-Verursachung ist – auch in diesem Fall nicht ersichtlich was als kausale Pairing Beziehung P fungieren könnte. In anderen Worten gilt es herauszufinden welche Beziehung zwischen Franz' und Franziskas Seele bzw. zwischen Fritz' und Friedas Seele besteht, die nicht auch zugleich Franz' und Friedas bzw. Fritz' und Franziskas Seele verbindet.

Zunächst möchte man davon ausgehen, dass es sich hierbei um irgendeine psychologische Beziehung handelt – doch wie sollte sie aussehen? Es müsste wohl es eine Art intentionale Beziehung wie ‚denken an‘, ‚herausgreifen‘ oder ‚sich beziehen auf‘ sein? Irgendwie müsste Franz' Seele schließlich Franziskas Seele identifizieren, um mit ihr in Kontakt zu treten. Wie dem auch sei, sollte etwa Franz' Seele an Franziskas Seele denken, sie herausgreifen oder sich auf sie beziehen, dann müssen die beiden Seelen, ähnlich einem Objekt außerhalb unseres Körpers, in irgendeiner kognitiven Relation zueinander stehen: diese kognitive Relation ist bei Objekten in der Außenwelt in der Regel eine Wahrnehmungsbeziehung – denn ein externes Objekt muss zunächst einmal wahrgenommen und von anderen es umgebenden Objekten unterschieden werden. Was aber bedeutet es, wahrgenommen zu werden? Um dies sinnvoll zu beschreiben muss ein kausales Element in der Erklärung vorhanden sein, wonach das Objekt die entsprechende Wahrnehmung verursacht. Letzten Endes scheint es so als ließen sich die intentionalen Beziehungen also nur auf Basis einer *kausalen* Beziehung erklären. Wenn dem aber so ist, dann können intentionale Beziehungen sicherlich nicht die Pairing Relation P sein, die Franz' und Franziskas Seele kausal vereinigt. Denn dies wäre eine klassische *petitio principii*: die intentionalen Beziehung zwischen den beiden Seelen würden erklären warum Franz' Seele mit Franziskas Seele in einer Kausalbeziehung zueinander stehen (und nicht etwa Franz' Seele mit Friedas Seele). Um aber die intentionalen Beziehungen zwischen dem Seelenpaar Franz-Franziska zu verstehen, müssten Kausalrelationen herangezogen werden.⁶⁶

Ein anderer Ansatz könnte zum Beispiel sein eine Art mentalen Raum einzuführen. In einem solchen Raum könnte jeder Seele zu einem bestimmten Zeitpunkt ein eindeutiger Ort zuge-

Außerdem ist davon auszugehen, dass es auch in interaktionistischen Dualismus eine Möglichkeit gibt wie Seele miteinander kommunizieren können; denn ein Dualist würde sicherlich zugestehen, dass etwa der Austausch von Argumenten ein mentaler Vorgang ist. Daher ist anzunehmen, dass die Schwierigkeit die Kommunikation zwischen zwei Seelen auszudrücken mit einem grundlegenden Problem des interaktionistischen Dualismus zu tun hat (nämlich mit der Nicht-Intelligibilität des Mentalen, vgl. oben Colin McGinn).

⁶⁶ Vgl. Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 80f.

wiesen werden. Damit würde auch für jede Paarung mentaler Entitäten, die miteinander wechselwirken, eine eindeutige ‚räumliche‘ Relation existieren, die über die Lokalisationen der Seelen im Mentalraum gegeben wäre. Auf diese Weise wäre auch eine Pairing Beziehung P sichergestellt, die erklärt warum ein bestimmtes Paar mentaler Entitäten kausal miteinander interagiert. Allerdings ist es völlig unklar welche psychologischen Relationen eine solche raum-ähnliche Struktur erzeugen könnten. Zudem ist bislang nichts dem Konzept eines Mentalraumes auch nur annähernd Ähnliches bekannt.⁶⁷

Wäre das Einführen des Begriffs der Kausalkette vielleicht eine Lösung für das Pairing Problem zwischen zwei mentalen Entitäten? Nach Hume bedürfen zwei kausal verbundene Ereignisse, die raumzeitlich getrennt sind einer Verbindung durch räumlich kontingente Ereignisse. Im Falle der Geist-zu-Geist-Verursachung ist es nun schwer ersichtlich wie eine solche Verbindung aussehen könnte. Eventuell sollte man eine zwischen A und A* vermittelnde dritte Seele C einführen. Dabei verursacht die Handlung der Seele A eine Veränderung in Seele C, die wiederum zum Ergebnis hat, dass in A* eine Veränderung verursacht wird. Allerdings ist auch diese Argumentation aus zwei Gründen nicht stichhaltig: erstens ist es gänzlich schleierhaft was das Wort ‚zwischen‘ in diesen Zusammenhang bedeuten soll. Schließlich sind A und A* per Definition nicht-räumlich, also kann das ‚zwischen‘ keine räumliche Bedeutung haben. Ist die Bedeutung hingegen kausal zu verstehen, kann das wiederum nicht die Pairing Beziehung P erklären (zumindest nicht auf eine nicht-zirkuläre Art und Weise) – schließlich ist eben genau von Interesse warum A und A* *kausal* miteinander verknüpft sind. Daher ist es tautologisch anzuführen A und A* seien kausal vereint durch eine Kausalrelation – dies führt zu keinerlei Erkenntnisgewinn.⁶⁸ Der zweite Grund warum es wenig aussichtsreich ist mit einer dritten Seele C eine vermittelnde Instanz zwischen A und A* einzuführen besteht darin, dass damit das Pairing Problem lediglich auf eine andere Ebene verschoben würde. Denn selbst wenn es gelingen sollte auf diese Weise eine adäquate Erklärung für die Pairing Beziehung P zwischen A und A* zu liefern, so würde sich unmittelbar die Frage anschließen welche Pairing Beziehung zwischen A und C (und analog zwischen C und A*) besteht. Das Einführen einer vierten Seele D, die nun zwischen A und C ‚geschaltet‘ ist, wäre sicherlich nicht sehr aussichtsreich. Schließlich ließe sich das Einführen von weiteren vermittelnden Seelen ad infinitum wiederholen. Ein solcher unendlicher Regress würde diesen Ver-

⁶⁷ Vgl. Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 82.

⁶⁸ Vgl. Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 82f.

such der Etablierung einer Pairing Beziehung P zwischen zwei mentalen Entitäten gewiß dis-
kreditieren.

Der Dreh- und Angelpunkt des Pairing Problems scheint genau in der Nicht-Räumlichkeit des Mentalen zu liegen. Denn eben diese Charaktereigenschaft des Mentalen scheint eine kausale Pairing Beziehung P – sei es zwischen einer mentalen Entität und einer physikalischen Entität oder zwischen zwei mentalen Entitäten – zu verhindern.

Unter der Prämisse, dass es in der Welt (mindestens) zwei grundverschiedene und nicht aufeinander reduzierbare Formen von Kausalität gibt – und zwar physikalische Verursachung und mentale Kausalität, stellt sich die Frage wie die Pairing Beziehung P zwischen zwei Ereignissen aussieht, wenn mentale Entitäten bei der Verursachung im Spiel sind. Bei Kausalbeziehungen zwischen physikalischen Gegenständen übernehmen *räumliche Beziehungen* zwischen den einzelnen Entitäten die Aufgabe der Pairing Beziehung P. Fällt etwa ein Dominostein und bringt dadurch den nächsten Stein in Reihe zum Umstürzen, so sind es allem voran die räumlichen Beziehungen der beiden Dominosteine (also ihre Anordnung im Raum und ihre räumliches Verhältnis zueinander), die das Umfallen des ersten Dominosteins mit dem anschließenden Fallen des zweiten Steins kausal in Verbindung bringen. Stehen nun mehr als zwei Dominosteine in einer Reihe, so lässt sich hinsichtlich des Fallens der Steine eine lückenlose Kausalkette vom letzten Stein bis hin zum Ersten rekonstruieren. Bei Kausalitätsbeziehungen, bei denen mentale Entitäten mitwirken, ist es hingegen alles andere als klar was in diesen Fällen als Pairing Relation P dienen sollte. Da räumliche Beziehungen zwischen mentalen Entitäten, aufgrund von deren nicht-räumlicher Natur, nicht in Frage kommen, müssen es folglich andere Beziehungen sein, die diese Rolle übernehmen. Es ist jedoch vollkommen nebulös wie solche Beziehungen stattdessen auszusehen hätten. Ähnlich fasst Kim seine Kritik am kartesischen Substanzdualismus zusammen:

(...) [O]ur idea of causation requires that the causally connected items be situated in a space-like framework. It has been widely believed (...) that Cartesian dualism of two substances runs into insurmountable difficulties in explaining the possibility of causal relations across the two domains, mind to body, and body to mind – especially, the former. But what our considerations show is that the problem runs deeper: the very same difficulties beset substantival dualism in regard to the possibility of mental-to-mental causation. Under substance dualism, mind-to-mind causation is no more intelligible than mind-to-body causation. Furthermore, the difficulty is rooted deep in the nature of immaterial minds: it is their supposed essential non-spatiality that makes it impossible for them to meet a basic requirement of causality, namely, the need for pairing relations. (...) A purely Cartesian world seems like a pretty lonely place, inhabited by immaterial souls each of which is an island unto itself, totally isolated from all other souls.⁶⁹

⁶⁹ Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 84f.

Alles in allem bringt das Pairing Problem, laut Kim, ein grundlegendes Manko des interaktionistischen Dualismus zum Ausdruck und stellt somit andere Probleme dieses Theoriegebäudes (wie zum Beispiel die Bedenken hinsichtlich der Intelligibilität des Wies mentaler Verursachung) in den Schatten, so schreibt er: „ (...) there is no need to appeal to the alleged “mechanical” nature of material causation and the supposed teleological or rational character of mental causation to show that mind-body causation is problematic. An effective argument can be formulated at a more general and basic level.“⁷⁰

Kritik am und Lösungen für das Pairing Problem

Schwache vs. starke Nicht-Intelligibilität

David Jehle zufolge zielt Kims Beschreibung des Pairing Problems darauf ab, dass die kausale Interaktion zwischen Körper und Seele im Substanzdualismus nicht intelligibel ist. Der Begriff der Intelligibilität ist nach Jehle aber doppeldeutig und wird von Kim nicht weiter erläutert. Jehle unterscheidet eine schwache Form der Nicht-Intelligibilität und eine starke Form. Nach der schwachen Version ist eine Position P genau dann schwach nicht intelligibel, wenn derzeit kein Mensch P versteht. Die starke Variante der Nicht-Intelligibilität aber besagt, dass eine Position P genau dann stark nicht intelligibel ist, wenn es unmöglich ist P überhaupt zu verstehen.

Für den Fall, dass Kim nun die starke Version der Nicht-Intelligibilität vertritt, lässt sich sein Argument, Jehle zufolge, wie folgt zusammenfassen:

- (1) If it is possible for souls to interact causally with material substances, then there must be *spatial* causal chains or *spatial* pairing relations between souls and material substances.
- (2) It is not possible that there are spatial causal chains or spatial pairing relations between souls and material substances.
- (3) Therefore, it is not possible for souls to interact causally with material substances.
- (4) And if it is not possible for souls to interact causally with material substances, then substance dualism is unintelligible.
- (5) Therefore, substance dualism is unintelligible.⁷¹

Zwar ist dieses Argument gültig, es beinhaltet laut Jehle jedoch zwei Schwierigkeiten: Erstens ist der modale Status von (3) alles andere als klar: Wenn die Modalität von (3) metaphysisch oder logisch verstanden würde, dann ließe sich die ontologische Unmöglichkeit von kausaler Wechselwirkung zwischen Seele und Körper beweisen. Allerdings, so behauptet Jehle, leug-

⁷⁰ Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 87f.

⁷¹ Jehle, David: „*Kim against Dualism*“ in: *Philosophical Studies*, Volume 130, 2006, S. 568f.

net Kim diese Möglichkeit.⁷² Folglich bleibe nur noch die Möglichkeit, dass es sich in (3) um eine nomologische Modalität handelt. Wenn das aber der Fall ist, dann ist (4) aber schlichtweg falsch. Denn es kann etwa nomologisch für mich unmöglich sein, dass ich in meinem Büro schwebe, das ist jedoch nicht gleichbedeutend damit, dass diese Vorstellung nicht intelligibel ist. Die zweite Schwierigkeit besteht darin, dass die Prämisse (1) eigentlich schon die Schlussfolgerung des Arguments enthält. Schließlich wird auf diese Weise implizit eine fundamentale Prämisse des Substanzdualismus, nämlich dass Seelen in ihrer Natur nicht-räumlich sind, von vornherein negiert.

Für den Fall, dass Kim die schwache Version der Nicht-Intelligibilität in seinem Argument meint, fasst Jehle das Argument so zusammen: „No one currently knows how souls and bodies are causally paired or how there would be causal chains between them. If so, then no one currently understands substance dualism. And if no one currently understands substance dualism, then substance dualism is not a useful or viable theory of mind.“⁷³ Es ist offensichtlich, dass dieses Argument falsch ist. Nach derselben Schlussregel ließe sich etwa auch beweisen, dass materialistische Theorien des Geistes keine gültigen Theorien sind. Schließlich weiß momentan niemand wie vor dem Hintergrund von materialistischen Theorien des Geistes phänomenales Bewusstsein entstehen kann.

Jehle folgert daraus, dass Kims Argument weder in der schwachen Lesart von Nicht-Intelligibilität noch in der starken Version (und somit überhaupt keinerlei) Überzeugungskraft besitzt.

Dispositionale Eigenschaften als Lösung des Pairing Problems

Neben seiner Kritik an Kims Argument formuliert Jehle auch einen Vorschlag wie ein Substanzdualist auf Kims Herausforderung reagieren könnte. Diese Herausforderung fasst Jehle wie folgt zusammen und skizziert anschließend einen möglichen Lösungsweg für den interaktionistischen Substanzdualismus:

Recall the challenge: explain why it is that, at *t*, when souls *a* and *b* simultaneously act, only *a* brings about a change in material substance *m*. The basic challenge here is for the substance dualist to explain how a soul is paired causally to one and only one body. This challenge is especially difficult for the substance dualist to meet says Kim, because she does not have spatial relations available to structure selective soul to body interaction. Actually, Kim seems to think that his challenge cannot be met. (...)

⁷² Jehle verweist in seinem Aufsatz in Fußnote 11 auf eine Stelle, an der Kim angeblich zu diesem Schluss gelangt. Nach genauem Studium der von Jehle angegebenen Textstelle, teile ich jedoch nicht seine Ansicht, wonach Kim an dieser Stelle, die ihm von Jehle unterstellte Ansicht vertritt. Mein Ziel besteht an dieser Stelle jedoch lediglich darin eine möglichst detailgetreue Darstellung von Jehles Kritik an Kim liefern.

⁷³ Jehle, David: „*Kim against Dualism*“ in: *Philosophical Studies*, Volume 130, 2006, S. 572.

So to meet Kim's challenge, the dualist needs to offer an possible explanation for how something like the following principle could be true.

(KC) $\Diamond \exists x \exists y [(x \text{ is a soul} \ \& \ y \text{ is a body} \ \& \ x \text{ is paired with } y) \ \& \ \forall z (z \text{ is paired with } x \rightarrow z = y)]$

In other words: the dualist needs to explain how it is possible that given a specific soul-body union, that no other body be paired with that soul and no other soul be paired with another body. If a substance dualist can provide some possible way to account for (KC), then there is, *pace* Kim, a way to render the idea of singular paired causation between a soul and body intelligible.⁷⁴

Jehle argumentiert weiter, dass es möglich ist (KC) Sinn zu verleihen, indem man eine bestimmte Auffassung von Kausalität vertritt. Danach ist Kausalität dadurch gekennzeichnet, dass Objekte bestimmte Eigenschaften besitzen – diese Eigenschaften sind Kräfte oder Dispositionen, die Objekten bestimmte kausale Kräfte übertragen. Folglich ist eine Kausalrelation eine gegenseitige Manifestierung von reziproken dispositionalen Partnern. Dies lässt sich an folgendem Beispiel recht anschaulich verdeutlichen: ein Salzkristall löst sich in einem Tropfen Wasser auf. Hierbei ist das Auflösen des Salzkristalls eine gegenseitige Manifestierung der Kraft des Wassers Salz aufzulösen (und der Kraft des Salzes in Wasser aufgelöst zu werden).⁷⁵ Vertritt man diese Auffassung von Kausalität, dann ließe sich laut Jehle folgende korrekte Ausformulierung von (KC) liefern:

(P) Some soul is such that it possesses a unique and singular property, such that in virtue of its possession of this property, it is able to causally interact directly with one and only one material substance by virtue of a property the material substance possesses and that property (of the material substance) also only allows for interaction with the particular soul in question.

The idea behind this account of soul-body pairing follows the account of causation offered above. Some soul *x* possesses a property and so also does some material substance *y*. These properties are such that both *x* and *y* are able to enter into a causal relationship because of the properties they possess allow for them to be specific dispositional partners. This account allows us to show that it is at least possible how a soul and body form a restricted causal union: namely, they are specific dispositional partners, paired by a property relation – a property relation that only holds between them and no other soul-body unit.⁷⁶

Schließlich führt Jehle selbst mehrere Kritikpunkte an, die gegen seinen Ansatz ins Feld geführt werden könnten. Die erste Strategie Jehles Theorie zu kritisieren, könnte an den von ihm angeführten Eigenschaften ansetzen. Man könnte behaupten diese Eigenschaften seien gänzlich unverständlich und verschleiern lediglich das Problem. Daher sei auch (P) nutzlos und müsse zurückgewiesen werden. Jehle selbst erwidert auf diese Kritik, dass sie im Wesentlichen metaphysische und epistemische Elemente vermenge: nur weil man nicht wisse welche Wahrheitsbedingungen (P) erfüllen muss, heißt das nicht, dass es unvernünftig ist (P) anzunehmen.⁷⁷ Zweitens könnte man gegen Jehle argumentieren (P) zwingt uns dazu zuzulassen,

⁷⁴ Jehle, David: „Kim against Dualism“ in: *Philosophical Studies*, Volume 130, 2006, S. 573.

⁷⁵ Vgl. Jehle, David: „Kim against Dualism“ in: *Philosophical Studies*, Volume 130, 2006, S. 573f.

⁷⁶ Jehle, David: „Kim against Dualism“ in: *Philosophical Studies*, Volume 130, 2006, S. 574.

⁷⁷ Vgl. Jehle, David: „Kim against Dualism“ in: *Philosophical Studies*, Volume 130, 2006, S. 574f.

dass nicht-allgemeine Gesetze einzelne Seelen mit einzelnen Körper verbinden. Aber trotz (P) gibt es gute Gründe anzunehmen, dass kausale Dispositionen und Kräfte von Objekten *allgemein* sind, dass sie also bestimmte Neigungen sind auf gewisse *Typen* von Objekten zu reagieren und nicht auf einzelne Objekte. Darauf erwidert Jehle, dass man diese bittere Pille schlucken muss und die Verbindung von einzelnen Seelen mit einzelnen Körpern durch nicht-allgemeine Gesetze zulassen müsse. Zudem verliere dieser Einwand an Attraktivität, wenn man bedenkt, dass Jehles Argument lediglich einen Gegenbeweis zu Kims Behauptung einer kausalen Verbindung zwischen einer bestimmten Seele und einem bestimmten Körper sei nicht intelligibel, darstellt. Schließlich sei (P) der Beweis dafür, dass dies eben doch kohärent denkbar sei. Ein dritter möglicher Kritikpunkt, den Jehle gegen seinen eigenen Ansatz formuliert, lautet: Wenn der Vorschlag Jehles stimme, dann wird eine materielle Eigenschaft P einer materiellen Substanz y mit einer immateriellen Eigenschaft P* einer immateriellen Substanz bzw. Seele x verknüpft. Dies erkläre jedoch nichts, sondern sage uns nur, dass Seelen und Körper durch eine Eigenschaftsrelation eineindeutig einander zugeordnet sind. Aber auch dieser Kritikpunkt gehe am Wesentlichen vorbei, so Jehle. Schließlich komme es nicht darauf an zu erklären, *wie* der Mechanismus für eine eineindeutige Interaktion zwischen einer Seele und einem Körper funktioniert – es genüge vielmehr zu zeigen, *dass* ein solcher Mechanismus vorstellbar ist. Schließlich bedeute, dass ϕ möglich ist, auch nicht dasselbe wie zu erklären wie ϕ funktioniert. Mit (P) habe Jehle lediglich den Gegenbeweis zu Kims Behauptung der Nicht-Intelligibilität einer kausalen Verknüpfung einer Seele mit einem Körper geliefert. Mehr könne man in diesem Zusammenhang aber auch nicht von (P) verlangen. Schließlich wolle Jehle nicht für die Wahrheit von (P) argumentieren, sondern nur aufzeigen, dass (P) ein möglicher Weg ist zu erklären, dass eine Seele und eine materielle Substanz kausal miteinander verknüpft sind.⁷⁸ Am Ende seines Aufsatzes resümiert Jehle seine Kritik an Kims Pairing Problem wie folgt: „The upshot: Kim’s challenge to the substance dualist to explain how a soul is paired with a body has been met. Substance dualism is widely dismissed because of the problem of interaction. Rarely does this dismissal come with a rigorous argument. Jaegwon Kim has attempted to provide just such an argument. We saw that the argument failed on all possible reconstructions. Nothing said here precludes a sensible rejection of substance dualism. But if substance dualism is to be rejected, it must be for reasons other than the ones Kim offers.“⁷⁹

⁷⁸ Vgl. Jehle, David: „*Kim against Dualism*“ in: *Philosophical Studies*, Volume 130, 2006, S. 574ff.

⁷⁹ Jehle, David: „*Kim against Dualism*“ in: *Philosophical Studies*, Volume 130, 2006, S. 575f.

Seelenfeld oder Mentalraum

Ein weiterer Kritiker von Kims Pairing Problem ist Timothy O'Connor. Sein Aufsatz ‚*Causality, Mind, and Free Will*‘ ist eine direkte Antwort auf Kims Aufsatz ‚*Lonely Souls*‘⁸⁰, in dem Kim das Pairing Problem darstellt. Am Anfang seines Aufsatzes fasst O'Connor Kims Argumentation hinsichtlich des Pairing Problems so zusammen:

(...) we can imagine two nonphysical minds with identical intrinsic states at time *t*0. Yet one acts on body B1 and the other acts on body B2. Why? Given that the minds bear no spatial relationships with these bodies, we must find some other kind of external relationship that explains the causal selectivity. Kim can think of none and concludes that there probably couldn't be one. So if causation cannot be haecceitistic, selective dualist interaction (one mind with one body, and vice versa) appears to be impossible. (...) Kim goes further. He suggests that the same consideration should lead us to suppose that causal interaction even *among* nonphysical minds is probably incoherent. For what kind of relationship could play the role of a structuring environment that is played by space for physical objects? If none is proffered, and we embrace the Eleatic linkage of existence with causal powers, we should suspect the coherence of the very idea of a nonphysical mind altogether.⁸¹

Unter ‚nicht-haecceitistischer‘ Kausalität versteht O'Connor eine Auffassung von Kausalität, wonach Objekte nicht eine einfache Disposition haben auf bestimmte andere einzelne Objekte zu wirken, sondern vielmehr die Disposition besitzen auf alle Objekte, die die geeigneten Eigenschaften haben, gleichermaßen zu wirken. Somit hat dasselbe Objekt zu unterschiedlichen Zeitpunkten dieselbe Wirkung auf verschiedene Tokens, die dem selben allgemeinen Typ angehören.⁸²

Zur Verteidigung der Mental-zu-mental-Verursachung im interaktionistischen Dualismus führt O'Connor zunächst ein ‚Seelensystem‘ ein. Angenommen Gott hätte eine Reihe von Seelen erschaffen und festgelegt, dass sie in ihrem Anfangszustand einem mathematisch beschreibbaren Feld gleichen, wobei jede einzelne Seele die Information bezüglich ihres aktuellen Standorts im Feld in Form eines primitiven, intentionalen Zustands mit sich trägt. Dem entsprechend ‚bewegen‘ sich Seelen in der Zeit durch das Feld, indem sie bestimmte Intentionen herausbilden und somit bestimmte Orte im Seelenfeld einnehmen. In diesem Seelensystem gibt es zwei verschiedene Arten kausaler Kapazität: einerseits die Möglichkeit bestimm-

⁸⁰ Beide Aufsätze sind auch in, der von Kevin Corcoran herausgegebenen Anthologie *Soul, Body, and Survival* erschienen. In dem dort veröffentlichten Aufsatz ‚*Lonely Souls*‘ argumentiert Kim wie in seiner später erschienenen (und von mir oben zitierten) Monographie *Physicalism or something near enough*. Da die besagte Monographie aber später erschienen und das Pairing Problem dort auch ausführlicher dargestellt ist, stammen die obigen Zitate daraus. Zudem geht Kim in dem von mir zitierten Buch auch bereits auf die Kritik von O'Connor geäußerte ein (s.o.).

⁸¹ O'Connor, Timothy: „*Causality, Mind, and Free Will*“ in: Corcoran, Kevin (Hg.): *Soul, Body, and Survival: Essays on the Metaphysics of Human Persons*, Ithaca, Cornell University Press, 2001, S. 45f.

⁸² Diese These der kausalen Allgemeinheit richtet sich im Wesentlichen gegen Singularisten, die behaupten Kausalität sei vor allem eine Beziehung zwischen *einzelnen* Ereignissen, die keinerlei Folgen dafür hat was zu einem anderen Zeitpunkt anderswo geschieht.

te Intentionen bezüglich sich selbst oder anderer herauszubilden und andererseits die Möglichkeit den intentionalen Zustand einer anderen Seele mittels der eigenen Intentionen zu ändern. Zur Erläuterung seines Gedanken nennt O'Connor folgendes Beispiel: „(...) by intending to communicate to Jaegwon the thought that it would be nice to have a body as humans do, I cause him to register this thought, along with a belief that it is *my* thought that he now is entertaining. Again, there will be some sort of dynamical law that governs the degree of success in bringing about such states in others: perhaps it will be directly proportional to the recipient's attentiveness and inversely proportional to his present informational load and “distance” in the array.“⁸³

Nun könnte man natürlich einwenden O'Connor verwende räumliche Metaphern um die Strukturen der Interaktion zwischen Seelen in einer qua Definition nicht-räumlichen Welt zu beschreiben. Dies sollte aber, O'Connor zufolge, kein Problem für sein Argument darstellen – schließlich würde auch in anderen Bereichen im übertragenen Sinne von ‚Raum‘ gesprochen, so etwa wenn in der Quantenmechanik vom ‚Phasenraum‘ die Rede ist, etc. Allerdings ließe sich aus diesem Bedenken auch folgender Vorwurf ableiten: „You have not actually described a framework of objective external relations. Instead, you have merely gestured at an abstract formalism and asserted without argument that there could be a kind of external relation so characterizable that is distinct from spatial relations and holds among nonphysical minds. Your reference to divine decree in instituting the array seems ineliminable, and in consequence the changing sequence over time that you described smacks of occasionalism, rather than real interaction.“⁸⁴ Nach O'Connor beraubt eine abstrakte, mathematische Charakterisierung einer objektiven Ordnung jene nicht des von ihr beschriebenen qualitativen Charakters. Zudem sind die einzigen externen Relationen zwischen Objekten, die man direkt verstehen kann, räumlicher oder zeitlicher Natur (abgesehen vielleicht vom schwierigen Fall der Kausalrelationen). Ferner gibt es auch andere Einsatzmöglichkeiten für solche mathematischen Strukturen: wenn man etwa daran glaubt, dass die Relativitätstheorie eine Richtschnur für die Metaphysik der Raumzeit darstellt, so wird man zugeben müssen, dass Raum und Zeit keine objektiven Größen sind und dass Raumzeitintervalle nicht unmittelbar erfahren werden können.

Für Skeptiker hinsichtlich der Möglichkeit externer Relationen in einem System von nicht-stofflichen Seelen, bestünde nach O'Connor eine andere Möglichkeit die Seele-zu-Seele-

⁸³ O'Connor, Timothy: „*Causality, Mind, and Free Will*“ in: Corcoran, Kevin (Hg.): *Soul, Body, and Survival: Essays on the Metaphysics of Human Persons*, Ithaca, Cornell University Press, 2001, S. 47.

⁸⁴ O'Connor, Timothy: „*Causality, Mind, and Free Will*“ in: Corcoran, Kevin (Hg.): *Soul, Body, and Survival: Essays on the Metaphysics of Human Persons*, Ithaca, Cornell University Press, 2001, S. 47f.

Interaktion zu retten. Und zwar darin, dass bei der Beschreibung des Systems auf externe Relationen verzichtet wird, indem jeder Seele ein primitiver intrinsischer Informationszustand zugeordnet wird, der bestimmt, wo sich die entsprechende Seele, in Beziehung zu allen anderen aufhält. Ob eine bestimmte Seele eine andere kausal beeinflusst hängt von ihrem Informationszustand und ihren anderen intrinsischen Eigenschaften ab. Somit sind alle kausalen Interaktionen gänzlich eine Funktion der intrinsischen Eigenschaften einer Seele und somit haben externe, strukturierende Relationen keinerlei Bedeutung.

Aber wie sieht es mit Geist-zu-Körper- und Körper-zu-Geist-Verursachung im kartesischen Dualismus aus – kann diese auch auf eine ähnliche Weise gerettet werden? Es sieht zunächst nicht so aus, so O'Connor:

We should require not just the system of ordered relations among the nonphysical souls, along with the system of spatial relations among the physical objects, but also a cross-grid mapping of the two, identifying in general terms which body will impact on which mind, and vice versa, in terms of their locations within their respective systems. And the trouble here is that on the Cartesian picture, we continually have the same pairings of individual souls and bodies, despite constant relational changes on at least the physical space side of the duality. The Cartesian picture of causal interaction seems unrepentantly haecceitistic.⁸⁵

Um eine konstante Konjunktion von bestimmten Seelen mit gewissen Körpern zu erreichen, bräuchte man also eine Abbildungsfunktion zwischen der Welt der nicht-stofflichen Seelen und der materiellen Welt der Körper. Es ist nun aber alles andere als klar ersichtlich wie eine solche quer zu beiden Rastern liegende Zuordnung stattfinden könnte und wie sie aussehen sollte.

Versucht man nun, ähnlich wie in obigem Fall, eine mathematische feldartige Struktur zu entwerfen, innerhalb welcher psychophysische Wechselwirkung stattfinden kann, steht man vor der Schwierigkeit, dass zumindest auf der Seite des physikalischen Raumes die Verortung der einzelnen Objekte nicht starr ist. Somit scheint es auch unmöglich zu sein eine feste Zuordnung zwischen einer bestimmten Seele und einem bestimmten Körper herzustellen.

Nun könnte man für ein zweidimensionales Geist-Körper-Feld argumentieren, das auf beiden Seiten externe Relationen erlaubt und zugleich nicht-räumlich ist. Allerdings dürften darin weder Seelen noch Körper ihre jeweilige Position verändern, so dass eine beständige kausale Verbindung einer Seele mit einem Körper möglich wäre. Nun verändern aber zumindest die Körper kontinuierlich ihre räumliche Beziehung zueinander, indem sie auf andere physikalische Objekte wirken und von solchen auch beeinflusst und somit bewegt werden. Angenommen man spräche den Seelen die Fähigkeit ab untereinander kausal zu interagieren, so wäre

⁸⁵ O'Connor, Timothy: „Causality, Mind, and Free Will“ in: Corcoran, Kevin (Hg.): *Soul, Body, and Survival: Essays on the Metaphysics of Human Persons*, Ithaca, Cornell University Press, 2001, S. 49.

doch vorstellbar, dass etwa Gott Seelen und Körper in dem zweidimensionalen Feld so anordnen könnte, dass sie auf Objekte der jeweils anderen Kategorie kausal wirken könnten. Kann man sich ein solches zweidimensionales Feld kohärent vorstellen? Wenn es in der realen Welt existieren würde, wäre man jedoch genötigt ein System physikalischer Relationen zu postulieren, das sich gänzlich der gewöhnlichen Betrachtung entzieht und das auch keinerlei Bedeutung für herkömmliche physikalische Kausalität (d.h. Körper-zu-Körper-Verursachung) hat. Da diese Auffassung aber mit gravierenden Konsequenzen für unser Weltbild behaftet ist, verfolgt O'Connor eine andere Strategie. Zunächst stellt er fest was für eine befriedigende Lösung des Pairing Problems von Nöten wäre: „What the dualist needs, as even Descartes saw but failed to provide, is a metaphysics on which mind and body constitute a unified natural system. We want a plausible picture on which a particular mind and body are not independent objects that somehow continually find one another in the crowd of similar such objects, but instead constitute a unified single system whose union is grounded independently of particular mental-physical interactions.“⁸⁶ Die Hauptschwierigkeit für den interaktionistischen Substanzdualismus sieht O'Connor also darin, eine kohärent vorstellbare konstante Konjunktion von einer Seele mit genau einem Körper zu begründen. Dieser Schwierigkeit war sich bereits Descartes bewusst und um dieses Problem zu umgehen postulierte er die Einheit von Körper und Geist als Axiom seiner Theorie des Geistes. Allerdings ist sein Rekurs auf eine notion primitive, also einen nicht weiter erklärbaren Grundbegriff des Denkens, nicht unbedingt zufriedenstellend.

Eine mögliche Art und Weise wie der interaktionistische Substanzdualismus das Pairing Problem lösen kann, sieht O'Connor in folgender Strategie:

Here is a way individual souls and bodies might constitute a single natural system. At some specific juncture in the development of the human organism, the body generates a soul, a nonphysical substance. Provided with the requisite degree of structural complexity and life-conserving functions of the body are preserved, the soul will likewise persist. Thus, it is completely dependent on the body not just for its coming to be but also for its continuing to be. Given such a baseline, asymmetrical dependency-of-existence relation, it is not arbitrary that these two entities should also interact continuously in more specific ways over time. On this picture, the soul is not entirely an entity in its own right, but is more properly seen as an aspect of the overall, fundamentally biological system that is the human person.⁸⁷

Diese substanzdualistische Form einer Emergenztheorie, wäre O'Connor zufolge, in der Lage Kims Kritik hinsichtlich des Pairing Problems zu überwinden. Allerdings wären auch die von ihr postulierten kausalen Kräfte des biologischen Systems außergewöhnlich groß: denn es

⁸⁶ O'Connor, Timothy: „Causality, Mind, and Free Will“ in: Corcoran, Kevin (Hg.): *Soul, Body, and Survival: Essays on the Metaphysics of Human Persons*, Ithaca, Cornell University Press, 2001, S. 49.

⁸⁷ O'Connor, Timothy: „Causality, Mind, and Free Will“ in: Corcoran, Kevin (Hg.): *Soul, Body, and Survival: Essays on the Metaphysics of Human Persons*, Ithaca, Cornell University Press, 2001, S. 50.

wäre möglich, dass aus der physikalischen Substanz eine völlig neue Substanz geschaffen wird – sozusagen eine creatio ex nihilo.⁸⁸

Es ist fraglich, ob diese Variante des interaktionistischen Substanzdualismus tatsächlich widerspruchsfrei gedacht werden kann. Denn naturgemäß würde sich die Frage anschließen, auf welche Art und Weise genau die geistige aus der körperlich Substanz entsteht (oder sich aus ihr herausbildet oder von ihr/sich geschaffen wird,...). Es ist jedoch zu erwarten, dass Vertreter dieser Form von Emergenztheorie eine Antwort auf diese Frage schuldig bleiben.

O'Connor denkt das Pairing Problem ließe sich besser lösen, wenn es auf einer abgeschwächten Variante des Dualismus basiert: „I suggest that those of us with dualist predilections try to live with a weaker form of dualism, on which token mental events are ontologically sui generis, distinct from any complex token physical state, without there being any substance distinct from the body which is the direct bearer of those events. This is a substance monism on which human persons are fundamentally biological entities that also have emergent mental states.“⁸⁹ Den Hauptgrund eine Eigenschafts- und keine Substanz-Emergenztheorie zu vertreten sieht O'Connor darin, dass auf diese Weise Körper und Seele als natürliche Einheit begriffen werden können (ohne dass dabei mysteriöse Kausalkräfte der physikalischen Substanz notwendig wären). So kommt O'Connor auch zu folgendem Fazit: „In summation, we have seen that property emergentism allows for a form of dualism that escapes Kim's problem, since the mental-physical interactions it posits occur within the context of a natural unitary system, and hence is a form of self-causality, rather than multiple-object transaction.“⁹⁰

Individualistische Kräfte

Ein weiterer Philosoph, der sich mit den Problemen des interaktionistischen Substanzdualismus befasst ist Peter Unger. Obwohl er sich in seinem Buch *All the Power in the World* nicht explizit auf Kims Pairing Problem bezieht, versucht er darin doch genau diesem Argument den Nährboden zu entziehen. Dabei beruft er sich auf individualistische Kräfte, die es Entitäten ermöglichen nur mit *bestimmten* Token eines Typs kausal wechselzuwirken und *nicht* mit

⁸⁸ Für eine kohärente Lösung von Kims Pairing Problem ist es, so O'Connor, nicht ausreichend auf Gedanken wie die prästabilisierte Harmonie oder das konkrete Eingreifen Gottes in jede kausale Interaktion zwischen Körper und Seele (i.e. Okkassionalismus) zurückzugreifen. Denn in beiden Fällen könnten Körper und Seele nicht als ein natürliches, einheitliches System verstanden werden.

⁸⁹ O'Connor, Timothy: „Causality, Mind, and Free Will“ in: Corcoran, Kevin (Hg.): *Soul, Body, and Survival: Essays on the Metaphysics of Human Persons*, Ithaca, Cornell University Press, 2001, S. 51.

⁹⁰ O'Connor, Timothy: „Causality, Mind, and Free Will“ in: Corcoran, Kevin (Hg.): *Soul, Body, and Survival: Essays on the Metaphysics of Human Persons*, Ithaca, Cornell University Press, 2001, S. 54.

anderen *intrinsisch identischen* Token desselben Typs. Die Idee der individualistischen Kräfte führt Unger mit folgendem Gedankenexperiment ein:

In a certain possible World, there may be a certain single Small Blue Particle that strongly attracts, and that's strongly attracted by, only a *certain single* Large Red Particle. This particular Small Blue, that we're so singularly imagining, *won't* strongly attract *any other entity at all*, other than the aforementioned Large Red, no matter how Large or how Red, or however nicely situated, any such other individual ever might be. And, for this part, this selfsame Large Red Particle may be strongly attracted by (and it may strongly attract) only this one Small Blue Particle, and not any other individual thing, no matter how like this Blue Sphere, in all *its Generalistic* features, any other concretum may be, and no matter how nicely situated may be any other (such terribly similar Small Blue) concrete particular.

Let's imagine, now, that our original Large Red is suddenly whisked far away, even into a galaxy that's many light-years distant. And, as we're imagining, in what was that Large Red's place just a moment ago, there is now *another* Large Red, precisely like the first in all Generalistic respects – it's precisely as Large, and it's precisely as Red, and so on. Will our selected Small Blue strongly attract, or be strongly attracted by, this precisely similar Johnny-come-lately? No, it *won't*; not in a World just as the one I've just specified. (...)

How shall we characterize how it is Propensitively with these peculiarly finicky particulars? By contrast with concrete particulars that are all just Generalistically Propensitied, we may say they're *Individualistically Propensitied* Individuals or, for short, concreta with *Individualistic Powers*.⁹¹

In diesem Abschnitt stellt Unger seine Idee von Einzeldingen mit individualistischen Kräften vor. Was versteht er unter diesen ominösen individualistischen Kräften? Er führt seine Idee anhand zweier miteinander wechselwirkender Elementarteilchen ein – und zwar ziehen sich in seinem Gedankenexperiment ein bestimmter Großer Roter Partikel und ein gewisses Kleines Blaues Teilchen gegenseitig sehr stark an. Die Besonderheit an den von Unger beschriebenen Teilchen besteht gerade darin, dass sie mit individualistischen Kräften ausgestattet sind. Aus diesem Grunde ziehen sich eben *nur dieses eine* Große Rote Elementarteilchen und *ausschließlich genau der besagte* Kleine Blaue Partikel gegenseitig an. Würde man nun das Große Rote Teilchen durch anderes Großes Rotes Teilchen *mit identischen intrinsischen Eigenschaften* austauschen, so würde damit in der von Unger beschriebenen möglichen Welt auch die *gegenseitige Anziehungskraft* der beiden Partikel *aufgehoben*.

Im Gegensatz dazu würden Vertreter eines *generalistischen* Ansatzes behaupten, dass der *Austausch* eines der beiden ursprünglichen Elementarteilchens mit einem *intrinsisch identischen* anderen Teilchen *keine Auswirkung auf die reziproke Anziehungskraft* der beiden Partikel hätte: Ersetzte man also das Große Rote Teilchen mit einem in jeder Beziehung identischen, anderen Großen Roten Elementarteilchen, so würde dieses nach der Vorstellung eines Vertreters generalistischen Gedankenguts genauso wie der erste Große Rote Partikel mit dem ursprünglichen Kleinen Blauen Teilchen wechselwirken – die beiden Teilchen würden sich also weiterhin gegenseitig anziehen. Kurzum würde jedes intrinsisch identische Token, das dem Typ ‚Großer Roter Partikel‘ angehört, mit dem ursprünglichen Kleinen Blauen Teilchen (kausal) wechselwirken; natürlich würde das auch für den umgekehrten Fall gelten – wenn

⁹¹ Unger, Peter K.: *All the Power in the World*, Oxford, Oxford University Press, 2005, S. 242f.

also der Kleine Blaue Partikel durch ein intrinsisch identisches Pendant substituiert werden würde.

Obwohl zwischen generalistischen und individualistischen Kräften offensichtlich ein großer Unterschied besteht, macht Unger dennoch einige Gemeinsamkeiten aus: „In *both* cases, the Propensity in question is intrinsic to the entity whose Propensity it is; the entity’s having the propensity doesn’t require, in either case, the existence of any other concreta. And, in *both* cases, there’s something *else*, in the *neighborhood* of the Propensity, that *does* require there to be an appropriate external entity, or entities. This is, of course, the manifestation of the Propensity in question, a *manifestation* that must involve, by hypothesis, the *interaction between* (or the *interaction among*) reciprocally Propensitized partners.“⁹²

Zunächst mag sich die Vorstellung von mit individualistischen Kräften ausgestatteten Einzel-
dingen obskur anhören – dessen ist sich auch Unger bewusst:

This is very different, I imagine, from how it is that, in almost all our thinking, we suppose matters to proceed in any physical realm, or reality. As our usual thinking goes, whatever particularity may be involved in the interaction among physical concreta, that will be determined by which particular places are the locations of which physical individuals, at which particular times, along with Generalistic features of the concreta in question, including their Generalistic Powers. Or, at the least, something to a similar effect will be in effect. Accordingly, the notion of Individualistically-directed Powers doesn’t figure greatly in our usual thinking, including our standard scientific thinking, at least not insofar as such thinking concerns only physical concreta.⁹³

Betrachtet man kausale Wechselwirkungen in der Welt, so scheint es als seien dort – zumindest im physikalischen Bereich – ausschließlich Dinge mit generalistischen Kräften am Werk.⁹⁴ Somit stellt sich natürlich die Frage, ob man die Möglichkeit individualistischer Kräfte überhaupt Ernst nehmen muss. Daher bittet Unger um einen kleinen Vorschuss an Vertrauen: „To folks gripped by our standard scientific thinking, myself so often included, any idea of Individualistically-directed Propensities may seem utterly incredible. Indeed, any such idea may seem so bizarre as to be worth no consideration. (...) My hope is that you’ll continue to give my work your thoughtful attention.“⁹⁵

⁹² Unger, Peter K.: *All the Power in the World*, Oxford, Oxford University Press, 2005, S. 244f.

⁹³ Unger, Peter K.: *All the Power in the World*, Oxford, Oxford University Press, 2005, S. 244.

⁹⁴ Tatsächlich würde es unser Weltbild beträchtlich erschüttern, wenn der (physikalischen) Kausalität dieser generalistische Aspekt abginge: würden sich zum selben Ereignistyp gehörige Einzelereignisse in Kausalrelationen nicht mehr genauso verhalten wie andere Tokens dieses Ereignistyps, wäre die Welt ein im wahrsten Wortsinne unberechenbarer Ort – empirische Wissenschaften wären (je nach Grad der Fluktuation) gänzlich unmöglich.

⁹⁵ Unger, Peter K.: *All the Power in the World*, Oxford, Oxford University Press, 2005, S. 246.

Den erbetenen Vertrauensvorschuss gedenkt Unger einzulösen, indem er mit Hilfe seiner Vorstellung von individualistischen Kräften eine kohärente Theorie des Geistes im Sinne Descartes' entwirft. Denn anhand der von ihm postulierten individualistischen Kräfte könnte sich der interaktionistische Dualismus einer bohrenden Frage entledigen – und zwar wie die konstante Verbindung einer bestimmten Seele mit einem bestimmten Körper zustande kommt bzw. überhaupt möglich ist. Diese Frage stellt sich auch Unger und beantwortet sie umgehend:

You ask: "How can it be that my body interacts with just *my* immaterial mind, and not also with (or even only with) the always very similar soul of Mr Duplicate, always so similar in all Generalistic respects, whose Generally (or Generalistically) so similar body is many galaxies removed from my body?" (It can't be that my immaterial mind is closer to my body, of course, since this nonspatial soul hasn't any spatial location at all.) Well, the answer lies in our idea of Individually-directed Propensities. Among all the World's many immaterial minds, your body is Propensitized with respect to only one mind that's your mind, and not with respect to your Duplicate's mind; whereas your Duplicate's body is Propensitized with respect to only *his* mind, and not with regard to yours. And, for its part, of course, your mind is Propensitized with respect to just your body, not any body but yours, while your Duplicate's mind is Propensitized for interaction with just his body.⁹⁶

Es ist offensichtlich welche Auswirkung die Möglichkeit von individualistischen Kräften auf das von Kim beschriebene Pairing Problem des interaktionistischen Substanzdualismus hätte: Wären nämlich eine bestimmte Seele und ein bestimmter Körper mit solchen individualistischen Kräften ausgestattet, so würde dies wiederum erklären warum nun gerade *diese* Seele mit eben *jenem* Körper in kausale Interaktion tritt und nicht mit einem intrinsisch identischen⁹⁷ Körper (und umgekehrt). Folglich würde dann Ungers individualistischen Kräften die Funktion der Pairing Relation zuteil – sie würden erklären warum eine bestimmter Körper mit einer gewissen Seele wechselwirkt und nicht mit einer in allen Eigenschaften identischen Kopie der Originalseele. In diesem Fall wären dann auch Körper und Seele nicht, wie von Kim behauptet, auf räumliche Relationen angewiesen, die – ähnlich wie im Falle physikalischer Kausalität – erklären könnten auf welche Art und Weise Ursache und Wirkung im Falle psychophysischer Wechselwirkung kausal gepaart sind.

Allerdings besteht eine weit verbreitete Aversion gegen individualistische Kräfte. Den Grund hierfür beschreibt Unger wie folgt: „Our antipathy to Individually-directed Propensities, which I myself find hard to overcome, shows how strongly we're gripped by the Hypothesis

⁹⁶ Unger, Peter K.: *All the Power in the World*, Oxford, Oxford University Press, 2005, S. 246f.

⁹⁷ Eigentlich geht Unger davon aus, dass weder Körper noch Seele von mir und von meinem Duplikat intrinsisch vollkommen identisch sein können: sie unterscheiden sich nämlich zumindest darin welche individualistischen Kräfte sie jeweils besitzen und da die individualistischen Kräfte per definitionem zum intrinsischen Aufbau meines Körpers bzw. meiner Seele gehören, kann es nicht der Fall sein, dass es eine absolute intrinsische Kopie meines Körpers bzw. meiner Seele gibt. Insofern ist auch die Redeweise von intrinsischen Dubletten meines Körpers bzw. meiner Seele widersinnig. So Unger: „But, then, typical philosophical talk of My Duplicate, or My Duplicate's body, will be as contentious as are any of the propositions that this talk presupposes.“ (Unger, Peter K.: *All the Power in the World*, Oxford, Oxford University Press, 2005, S. 247.)

that All Physical Propensities concern only General Features of Entities. Though it's implicitly accepted by us all, that substantive Hypothesis clearly isn't any purely *conceptual* truth. Indeed, it may be only with something like a *Scientiphical Faith*, I imagine, that we take our actual world to be a world for which the Hypothesis holds true."⁹⁸ Ungers exakte Definition besagter Hypothese lautet: „*The Hypothesis that All Physical Propensities concern only General Features of Entities*. Whenever a physical entity has a Propensity with regard to physical reality, the power always concerns *only some general feature, or features, of some physical things* (either some general features of such entities as are its disposition partners or, in a more special case, some general features of the physical entity itself."⁹⁹ Die Abneigung gegen individualistische Kräfte beruht Unger zufolge also auf der Hypothese, dass physikalische Einzeldinge nur aufgrund ihrer generalistischen Eigenschaften miteinander in Wechselwirkung treten. Es ist offensichtlich, dass sollte sich diese Hypothese als wahr behaupten, dem Glauben an individualistische Kräfte der Boden entzogen werden würde. Allerdings spiegelt die Hypothese laut Unger keine analytische Wahrheit wider, sondern sie ist vielmehr bestenfalls eine Art wissenschaftlicher Glaube. Es ist anzunehmen, dass Unger sogar wohl noch weiter gehen und die Hypothese als einen wissenschaftlichen *Aberglauben* brandmarken würde, den es zu überkommen gilt. Andererseits sieht er aber auch, dass die Hypothese tief ins Bewusstsein der Menschen eingegraben ist und dass diese Verinnerlichung nur mit großer (intellektueller) Anstrengung überkommen werden kann.

Die nomologische Annahme und personengebundene Gesetze

Nach John Foster beruht das Pairing Problem im Wesentlichen auf einer Annahme, die er als ‚nomologische Annahme‘ bezeichnet. Der ‚nomologischen Annahme‘ zufolge, kann das Auftreten einer Kausalbeziehung immer aus gewissen nicht-kausalen Merkmalen der jeweiligen Situation zusammen mit dem Wirken (eines oder mehrerer) Gesetze abgeleitet werden. Foster verdeutlicht dies an folgendem Beispiel:

It is commonly assumed that where two events are causally related, their being so is wholly constituted by the way in which, via their non-causal properties and relations, they fall under some natural law or set of laws. Thus if, on a particular occasion, my heating a lump of metal causes it to melt, it is assumed that what ultimately makes it true is that the metal is of a certain intrinsic type, that on the occasion in question it reaches a certain temperature, and that it is a law of nature (or a consequence of a law or set of laws) that whenever metal of that type reaches that temperature, it melts. And quite generally, it is assumed that whenever one event causes another, the obtaining of this causal relationship is derivative from, nothing over and above, certain non-causal

⁹⁸ Unger, Peter K.: *All the Power in the World*, Oxford, Oxford University Press, 2005, S. 248.

⁹⁹ Unger, Peter K.: *All the Power in the World*, Oxford, Oxford University Press, 2005, S. 240.

aspects of the situation and the obtaining of certain relevant covering laws. Let us refer to this as the ‘nomological assumption’.¹⁰⁰

Versucht man nun diese nomologische Annahme, vor dem Hintergrund einer dualistischen Ontologie, auf psychophysische Wechselwirkung anzuwenden steht man vor einem Problem. Denn sollte es ein gleichzeitiges Duplikat des mentalen oder des physikalischen Ereignisses geben, kann ein Dualist kein Gesetz angeben, das eine eindeutige Zuordnung der jeweiligen kausalen Paare liefert. In anderen Worten, man steht vor dem sogenannten Pairing Problem. Hierzu ein Beispiel: Angenommen ein neuronaler Zustand N, der zum Zeitpunkt t in Smiths Gehirn stattfindet, verursacht eine Zehntelsekunde später direkt eine Schmerzempfindung E in Smiths Geist. Welches Gesetz könnte ein Dualist nun angeben? Blicke er der nomologischen Annahme verhaftet, müsste sein Gesetz etwa so aussehen:

L1: It is a law that whenever an event of physical type Φ occurs in the brain of structural type Σ an experience of psychological type Ψ occurs a tenth of a second later¹⁰¹

Wobei gilt, dass N vom Typ Φ ist und Smiths Gehirn dem strukturellen Typ Σ angehört und schließlich E vom psychologischen Typ Ψ ist. Es ist jedoch augenscheinlich, dass sich ein Dualist dann genau mit dem Pairing Problem konfrontiert sieht. Denn ließe sich die Kausalbeziehung zwischen N und E durch L1 und die nicht-kausalen Eigenschaften der Situation erklären, dann käme an folgendem Punkt das Pairing Problem zum Tragen: Angenommen die kausale Abfolge in Smith würde dupliziert und fände gleichzeitig in einer anderen Person, sagen wir Jones, statt. Auch Jones’ Gehirn gehört zum Zeitpunkt t dem strukturellen Typ Σ an und das Ereignis N’ vom Typ Φ findet in Jones’ Gehirn zum Zeitpunkt t statt. Schließlich gehört auch die Schmerzempfindung E’ dem Typ Ψ an. Eigentlich sollte man nun annehmen können, dass N E und N’ E’ verursacht – aber diese kausalen Paare lassen sich eben *nicht* aus den nicht-kausalen Eigenschaften der Situation gemeinsam mit L1 ableiten. Worin liegt der Grund hierfür? Dazu konstatiert Foster: „The inadequacy of L1 stems from the fact that it only links the specified physical type of event and the specified psychological type of event by means of a temporal relation. For, of course, this relation cannot help to settle the causal pairings of *simultaneous* events of either type. The natural remedy would be to supplement the temporal relation with some further relation (...)“¹⁰² Doch welche Beziehung könnte die

¹⁰⁰ Foster, John: *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*, London, Routledge, 1991, S. 163.

¹⁰¹ Foster, John: *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*, London, Routledge, 1991, S. 164.

¹⁰² Foster, John: *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*, London, Routledge, 1991, S. 164.

zeitliche Relation ergänzen? Wie bereits mehrfach angesprochen kommen räumliche Beziehungen an dieser Stelle nicht in Frage – schließlich besitzen mentale Ereignisse, einer der Kernthesen des Substanzualismus zufolge, keine Lokalisierung im Raum. Nach Foster ist die Antwort auf diese Frage offensichtlich:

The obvious candidate is the relation which holds between any mental event in some subject's mind and any physical event in the *same* subject's body – in other words, the relation signified by the expression 'x belongs to a subject in whose body y occurs'. Let us, for convenience, abbreviate this expression to 'x is subject-linked to y'. Then the dualist could suppose that, in the case of Smith and Jones, there is a covering law which nomologically links the relevant types of event both temporally and under this additional relation, i.e.

L2: It is a law that, for any Φ -event x in a brain of type Σ , there is a Ψ -experience y such that y is a tenth of a second later than x and y is subject-linked to x .¹⁰³

Es ist offenkundig, dass L2 eindeutige kausale Paare für das oben angeführte Beispiel von Smith und Jones liefert. Schließlich ist die konstante Konjunktion eines bestimmten Körpers und des ‚ihm zugehörigen‘ Geistes ein integraler Bestandteil von L2.

Diese Ausweichroute gegenüber dem Pairing Problem setzt jedoch etwas Entscheidendes voraus – und zwar, dass die Einheit eines bestimmten Geist-Körper-Paares in diesem Zusammenhang völlig unproblematisch angenommen werden kann. Allerdings ist eben diese Einheit in ihrem Kern *kausaler* Natur, d.h. der Grund warum ein bestimmter Geist einem bestimmten Körper (oder umgekehrt) zugeordnet ist, besteht hauptsächlich darin, dass sie kausal aneinander gebunden sind: man erkennt die Einheit von Körper und Geist eben daran, dass sich exakt dieser Geist und genau dieser Körper gegenseitig kausal beeinflussen. Versucht man nun diese kausale Gebundenheit mit L2 zu erklären, resultiert daraus augenscheinlich ein zirkuläres Argument. Ist man hingegen bestrebt die kausale Verbindung anhand von L1 zu erklären, so hat das wiederum unklare kausale Zuordnungen zur Folge – es wäre nicht klar welcher Körper zum welchem Geist gehört.

Das Argument, dass daraus gegen den interaktionistischen Substanzdualismus geschmiedet wird, könnte man in etwa wie folgt zusammenfassen: erst wenn man anerkennt, dass mentale Ereignisse identisch¹⁰⁴ mit physikalischen Ereignissen sind, kann das Pairing Problem vermieden werden. Denn wenn mentale Ereignisse physikalisch sind, besitzen sie auch räumliche Ausdehnung und stehen somit in raumzeitlichen Beziehungen zu anderen (physikalischen) Ereignissen. Dies wiederum ermöglicht eine exakte Bestimmung der raumzeitlichen Anord-

¹⁰³ Foster, John: *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*, London, Routledge, 1991, S. 165.

¹⁰⁴ Das soll nicht heißen, dass zwischen mentalen und physikalischen Ereignissen einen *Typenidentität* bestehen muss: ein mentaler Ereignistyp wie etwa ‚Schmerz‘ müsste demzufolge nicht unbedingt mit einem physikalischen Ereignistyp wie ‚C-Faser-Feuern‘ identisch sein. Es reicht aus, wenn zumindest mentale und physikalische Einzelereignisse (tokens) miteinander identisch sind. Kurzum, es genügt eine *token identity theory* zu vertreten.

nung von Ursachen und deren Wirkungen zueinander. Infolgedessen lässt sich auch eine eindeutige kausale Pairing Beziehung angeben, die Ursache und Wirkung direkt miteinander verknüpft. Sollte dies alles richtig sein, dann hat Kim vermutlich Recht und die radikale Verschiedenheit von mentalen und physikalischen Phänomenen im Dualismus verhindert eine kohärente Erklärung der kausalen Wechselwirkung zwischen Mentalem und Physikalischem. Foster geht jedoch davon aus, dass dem interaktionistischen Dualisten vor diesem Hintergrund nicht die Hände gebunden sind, sondern dass das Pairing Problem sogar zwei Lösungen besitzt:

(...) I think that the dualist does have the resources to deal with the problem – and in two quite different ways. The argument against him rests on two premises. The first of these is the nomological assumption, that causal relationships between events are always constituted by certain non-causal properties of the situation, together with the relevant covering laws. The second is the claim that it is only by taking mental events to be physical that, in cases of duplication, we can envisage laws which cover the causal pairings in the way this assumption requires. Now it seems to me that the dualist is entitled to reject both these premises, the rejection of each constituting a separate solution to the problem.¹⁰⁵

Nach Foster beruht das Pairing Problem im Wesentlichen auf zwei Prämissen: die erste ist die oben beschriebene nomologische Annahme. Die zweite Voraussetzung für die Entstehung des Pairing Problems ist die Annahme, dass im Falle von intrinsischen Duplikaten mentale Ereignisse als physikalische Ereignisse verstanden werden müssen, um in Hinblick auf die nomologische Annahme eindeutige kausale Paare zu ermöglichen. Dementsprechend stehen einem Substanzdualisten zwei Wege offen das Pairing Problem aufzulösen. Die erste Variante zielt darauf ab, im Rahmen der nomologischen Annahme Gesetze zu beschreiben, die zu exakten kausalen Paarungen führen und gleichzeitig aber den dualistischen Grundannahmen treu bleiben. Der zweite Ansatz besteht darin die nomologische Annahme gänzlich zurückzuweisen.

Wie könnte es nun gelingen Gesetze zu finden, die einerseits der nomologischen Annahme gerecht werden, andererseits aber mit dem interaktionistischen Substanzdualismus vereinbar sind? Foster zufolge muss man sich hierzu eines einfachen ‚Tricks‘ bedienen. Dieser Kunstgriff besteht darin den Anwendungsbereich der Gesetze auf einzelne Personen zu beschränken. Im Falle von Smith und Jones könnte man sich das in etwa folgendermaßen denken:

- | | |
|-----------|---|
| L(Smith): | It is a law that whenever a Φ -event occurs in brain B^S at time t when B^S is of structural type Σ , a Ψ -experience occurs tenth of a second later in mind M^S |
| L(Jones) | It is a law that whenever a Φ -event occurs in brain B^J at time t when B^J is of structural type Σ , a Ψ -experience occurs tenth of a second later in mind M^J . |

¹⁰⁵ Foster, John: *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*, London, Routledge, 1991, S. 167.

where B^S and M^S are respectively Smith's brain and mind, and B^J and M^J are respectively Jones's brain and mind. These laws would secure the right causal pairings. N and E would be causally linked under $L(\text{Smith})$, and N' and E' would be causally linked under $L(\text{Jones})$.¹⁰⁶

Schränkt man den Geltungsbereich der Gesetze so wie Foster auf einzelne Personen ein, dann ist auch eine eindeutige Zuordnung zwischen Mentalem und Physikalischem gewährleistet. In anderen Worten kann das Pairing Problem für den Substanzdualisten als gelöst betrachtet werden. Allerdings werden Sympathisanten des Pairing Problems nicht so einfach einlenken und diesen Lösungsvorschlag akzeptieren. Im Gegenteil, es ist zu erwarten, dass sie dieses Argument sofort anfechten, indem sie den eingeschränkten Geltungsbereich der Gesetze monieren. Denn nach allgemeingültigem Verständnis zeichnen sich Gesetze gerade dadurch aus, dass sie einen *universellen* Geltungsbereich haben und ausschließlich in *allgemeinen* Begriffen beschrieben sind. Bei $L(\text{Smith})$ oder $L(\text{Jones})$ ist die offensichtlich nicht der Fall – denn sowohl ihr Geltungsbereich ist stark eingegrenzt (sogar nur auf eine einzige Person) und in ihrer Formulierung kommen spezielle Begriffe (eben Smith bzw. Jones) vor. Normalerweise ist man aber geneigt anzunehmen, dass Dinge, die unter ein bestimmtes Gesetz fallen, sich so oder so verhalten, weil sie einer bestimmten *Art* von Dingen angehören, unabhängig davon welches bestimmte Einzelding sie nun sind. Für Foster ist die Beschränkung des Geltungsbereiches der Gesetze in diesem Fall jedoch völlig unproblematisch, denn „[w]here there is a law for one person, there are exactly similar laws for everyone else: the only thing which varies is the reference to the particular brain and particular mind in question.“¹⁰⁷ Zudem lässt sich, so Foster, auch ein allgemeines und uneingeschränktes Gesetz formulieren, unter das die einzelnen Gesetze mit einem auf eine Person beschränkten Geltungsbereich subsumiert werden können – und zwar folgendes:

L3: It is a law that there is some 1-1 correlation between human brains and human minds such that any Φ -event in a brain X of structural type Σ is a tenth of a second earlier than some Ψ -experience in that mind which is correlated with X .¹⁰⁸

Zwar lassen sich die einzelnen Gesetze mit beschränkten Geltungsbereich nicht aus L3 ableiten, denn das Gesetz legt nicht fest wie Gehirne und Geister miteinander korreliert sind, aber L3 garantiert, dass es solche Korrelationen gibt. In Gesetzen deren Gültigkeit auf eine Person beschränkt ist, liegt also Fosters erste Lösungsmöglichkeit für das Pairing Problem. Aller-

¹⁰⁶ Foster, John: *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*, London, Routledge, 1991, S. 167f.

¹⁰⁷ Foster, John: *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*, London, Routledge, 1991, S. 168.

¹⁰⁸ Foster, John: *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*, London, Routledge, 1991, S. 168.

dings gibt es nach Foster eine Einschränkung für den umfassenden Erfolg dieses Lösungsansatzes – nämlich, dass ein Dualist dann nicht gleichzeitig die Ansicht vertreten kann, dass die Einheit des Geistes (ganz oder zum Teil) durch kausale Relationen von mentalen Entitäten zu demselben Körper oder Gehirn verbürgt wird. Ansonsten wäre die Theorie zirkulär: „(...) the account of what constitutes a single mind would presuppose the availability of a prior account of psychophysical causation, while, by postulating laws with irreducible references to particular minds, the account of psychophysical causation would presuppose the availability of a prior account of what constitutes a single mind.“¹⁰⁹

Der zweite Ansatz zur Lösung des Pairing Problems unterscheidet sich essentiell vom ersten, denn er versucht dem Problem die Grundlage zu entziehen – im engeren Sinne will er das Problem also nicht nur lösen, sondern vielmehr auflösen. Dieser Lösungsweg ist dadurch gekennzeichnet, dass er die sogenannte nomologische Annahme zurückweist. Wie oben bereits ausführlicher dargelegt, lässt sich die nomologische Annahme wie folgt zusammenfassen: Wenn zwei Ereignisse in einer Kausalrelation zueinander stehen, dann ist diese Verursachungsbeziehung vollständig durch die nicht-kausalen Eigenschaften der Situation und das Gelten von bestimmten Gesetzen festgelegt. Laut Foster kann nun auf zwei Arten gezeigt werden, dass die nomologische Annahme, obwohl sie weit verbreitet ist, falsch ist.

Erstens berücksichtigt die nomologische Annahme nicht die zeitliche Dimension von Verursachung. Es gäbe also keinen Grund rückwärts gerichtete Verursachung (also Fälle, in denen die Wirkung *zeitlich vor* der Ursache eintritt) abzulehnen. Somit steht die nomologische Annahme augenscheinlich im drastischen Gegensatz zu der Alltagsauffassung von Kausalität, wonach Verursachung ein zeitlich eindeutig gerichteter Prozess ist, bei dem ein Ereignis (die Wirkung) asymmetrisch von einem anderen Ereignis (der Ursache) abhängt. Dies verdeutlicht Foster anhand eines Beispiels:

(...) suppose one billiard ball strikes another, causing it to move. Let us assume that the covering law is such that, in the circumstances, the second type of event would not have occurred without the prior occurrence of the first, and the first could not have occurred without the subsequent occurrence of the second. The nomological relationship between the two types of events is the symmetric: the first is necessary and sufficient for the second, which is thereby sufficient and necessary for the first. Consequently, the nomological facts do not account for the directionality in causal processes: they do not explain the sense in which the first event is asymmetrically responsible for the second – the sense in which the second owes its occurrence to the first, but not vice versa. And this means that, contrary to the nomological assumption, their causal relationship transcends the obtaining of the relevant laws and the non-causal features of the situation.¹¹⁰

¹⁰⁹ Foster, John: *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*, London, Routledge, 1991, S. 169.

¹¹⁰ Foster, John: *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*, London, Routledge, 1991, S. 169f.

Da der nomologische Zusammenhang zwischen zwei Ereignissen symmetrisch ist, eignet er sich nicht zur Explikation von Kausalrelationen, die ihrem Wesen nach asymmetrisch sind. Wenn es aber für die adäquate Erklärung einer Kausalrelation mehr bedarf als die nicht-kausalen Merkmale einer Situation und die nomologischen Beziehungen zueinander zu kennen, dann ist dadurch bereits die nomologische Annahme falsifiziert.

Zweitens sind auch im physikalischen Bereich Fälle denkbar, bei denen (Natur-)Gesetze nicht die aktuellen kausalen Paarungen liefern. Auch für diesen Fall gibt Foster ein Beispiel an:

(...) suppose that there is a certain kind *K* of metal, and it is a law of nature that when any spherical *K*-lump reaches a specified temperature, a flash occurs tenth of a second later somewhere (unspecified) in the region of points which are no further from the centre of the sphere than twice its diameter. Suppose, further, that there is no stronger law to fix the location of the flash more precisely, and indeed that, at any moment when the critical temperature is reached, each position in the specified region has an equal chance of receiving it. Now we are surely entitled to construe the relationship between the occurrence of the critical temperature in a *K*-sphere and the subsequent occurrence of a flash in the specified region as a *causal* one: we are surely entitled to say that the temperature-event causes the flash-event. And normally, this interpretation creates no problems for the nomological assumption. But now imagine a case in which two *K*-spheres, which are sufficiently close together for their specified regions to overlap, reach the critical temperature simultaneously, and that, a tenth of a second later, two flashes occur within the region of overlap. It still seems right to suppose that each flash is the effect of just one of the temperature-events and that each temperature-event is the cause of just one of the flashes. But, because each flash falls within the specified region of each sphere, the causal pairings are not determined by the law and the non-causal conditions. So once again we have a situation in which the causal relationships transcend the factors which would wholly constitute them if the nomological assumption were true. Admittedly, the example is only fictitious; and for all I know, nothing like this occurs in the actual world. But its very conceivability is what matters in the present context. For unless the nomological assumption can be defended a priori, it cannot be invoked as part of an a priori objection to the dualist's position.¹¹¹

Mit diesem Gedankenexperiment versucht Foster zu zeigen, dass auch im physikalischen Bereich Fälle vorstellbar sind, bei denen unklar ist, was als kausale Pairing-Beziehung zwischen Ursache und Wirkung fungieren könnte: das gleichzeitige Erreichen der kritischen Temperatur in den beiden *K*-Punktwolken hat sicherlich das Auftreten der beiden Blitze verursacht. Es ist jedoch unmöglich anzugeben, welche *K*-Punktwolke das Auftreten welches Blitzes verursacht hat, weil beide Blitze in der Schnittmenge der beiden Punktwolken erschienen sind und kein Punkt in der Punktwolke eine höhere Wahrscheinlichkeit für das Auftreten eines Blitzes besessen hat. Da in diesem Fall keine eindeutige Zuordnung gemacht werden kann, fehlt auch an dieser Stelle eine Zuordnungsfunktion, die die jeweilige Ursache mit ihrer entsprechenden Wirkung kausal zusammenbringt. Somit sind Kausalbeziehungen auch im physikalischen Bereich (zumindest in einigen Sonderfällen) nicht vollständig durch alle nicht-kausalen Eigenschaften der Situation und die entsprechenden Gesetze festgelegt. Anders ausgedrückt, verliert die nomologische Annahme auch in der physikalischen Domäne ihre universelle Gel-

¹¹¹ Foster, John: *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*, London, Routledge, 1991, S. 170f.

tung. Folglich eignet sie sich nach Foster auch nicht als a priori Argument gegen den Dualismus.

Das Zurückweisen der nomologischen Annahme erweist sich nach Foster also als eine erfolgreiche Strategie das Pairing Problem zu lösen. Da diese Variante im Vergleich zu ersten Lösungsmöglichkeit (Gesetze mit beschränktem Geltungsbereich) potenter zu sein scheint, muss ein Substanzdualist nicht den Anwendungsbereich eines Gesetzes auf einzelne Personen einschränken, um so der nomologischen Annahme gerecht zu werden. Andererseits bedeutet das nicht, dass dieser Lösungsansatz völlig nutzlos ist. Denn während es bei den *K*-Punktwolken zunächst keine Rolle spielen mag welcher Temperaturanstieg welchen Blitz verursacht hat, ist die Sachlage bei psychophysischer Kausalität anders. Schließlich möchte ein Substanzdualist, dass die mentale Wirkung einer neuronalen Ursache (oder umgekehrt) eine Veränderung in *diesem einen* Geist (bzw. Gehirn) und nicht in irgendeinem anderen hervorruft. Die Interaktion zwischen Körper und Geist hängt also im Wesentlichen von einer psychophysischen Ordnung ab, die es garantiert, dass dieser Geist mit jenem Körper wechselwirkt. An diesem Punkt kommt nun der erste Lösungsansatz hinsichtlich des Pairing Problems ins Spiel: denn es erscheint sinnvoll, dass ein System von Gesetzen mit eingeschränktem Geltungsbereich die Funktion dieser psychophysischer Ordnung übernimmt. Durch ihren auf einen Geist und ein Gehirn begrenzten Anwendungsbereich wird der Einflussbereich beidseitig in einem für eine konstante Geist-Körper-Wechselwirkung erforderlichem Maße beschnitten.

Das Problem um die Anomalie des Mentalen

Eine sehr eindrückliche Beschreibung des Problems um die Anomalie des Mentalen und den wohl prominentesten Lösungsvorschlag dafür formuliert Donald Davidson 1970 in seinem epochalen Aufsatz *Mental Events*. Den Ausgangspunkt für Davidsons Essay bilden drei Prinzipien, deren Wahrheit für sich genommen jeweils als evident erscheint, die aber den Anschein erwecken als wären sie keinesfalls miteinander vereinbar. Genauer gesagt bringen diese drei Prinzipien laut Davidson einen vermeintlichen Widerspruch hinsichtlich mentaler Ereignisse oder in anderen Worten das Problem um die Anomalie des Mentalen zum Ausdruck. Die besagten drei Prinzipien beschreibt er wie folgt:

The first principle asserts that at least some mental events interact causally with physical events . (We would call this the Principle of Causal Interaction.)

(...)

The second principle is that where there is causality, there must be a law: events related as cause and effect fall under strict deterministic laws. (We may term this the Principle of the Nomological Character of Causality.)

(...)

The third principle is that there are no strict deterministic laws on the basis of which mental events can be predicted and explained (the Anomalism of the Mental).

The paradox I wish to discuss arises for someone who is inclined to accept these three assumptions or principles, and who thinks they are inconsistent with one another.

(...)

Many philosophers have accepted, with or without argument, the view that the three principles do lead to a contradiction. It seems to me, however, that all three principles are true, so that what must be done is to explain away the appearance of contradiction (...).¹¹²

Das erste Prinzip, das nach Davidson einen wesentlichen Anteil an der Entstehung des Problems um die Anomalie des Mentalen hat, nennt er das ‚Prinzip der kausalen Interaktion‘. Im Grunde besagt dieses Prinzip lediglich, dass es so etwas wie mentale Verursachung grundsätzlich gibt.

Prinzip der kausalen Interaktion (PKI): (Zumindest einige) Mentale Ereignisse interagieren kausal mit physikalischen Ereignissen.

Die These von der Existenz mentaler Kausalität scheint ein *factum brutum* der alltäglichen Erfahrung zu sein. Schließlich spielen mentale Ereignisse eine Schlüsselrolle bei der Erklärung menschlichen Verhaltens. So würde man gemeinhin beispielsweise sicherlich annehmen, dass zahlreiche mentale Ereignisse eine kausale Rolle beim Versenken des Schlachtschiffes Bismarck inne hatten: etwa in Form von Wahrnehmungen, Berechnungen, Urteilen, Entschei-

¹¹² Davidson, Donald: ‚*Mental Events*‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 208f.

dungen, Intentionen, usw. Genauer gesagt ist man gewiss geneigt davon auszugehen, dass wenn eine bestimmte Person den Untergang der Bismarck verursacht hat, letztendlich mentale Ereignisse dieser Person ursächlich für ihre Körperbewegungen waren, die am Ende dafür gesorgt haben, dass die Bismarck sank. Ein Paradebeispiel dafür wie Verursachung vom Physikalischen zum Mentalen verläuft sind freilich Wahrnehmungsprozesse: nimmt jemand nämlich beispielsweise wahr, dass sich ein Schiff nähert, dann muss das Herannahen des Schiffes den Glauben in der Person verursacht haben, dass ein Schiff auf sie zusteuert. Wenngleich auch Wahrnehmungen und Handlungen die offensichtlichsten Hinweise darauf geben, dass mentale Ereignisse mit physikalischen Ereignissen in kausale Wechselwirkung treten, so gibt es laut Davidson gute Gründe dafür anzunehmen, dass *alle* mentalen Ereignisse (ggf. vermittelt anderer mentaler Ereignisse) mit physikalischen Ereignissen kausal interagieren. Allerdings spiele dies weder für seine Darstellung noch für seine Lösung des Anomalie-Problems eine Rolle, weil PKI lediglich eine schwächere Behauptung aufstellt – und zwar, dass zumindest *einige* mentale Ereignisse physikalische Ereignisse verursachen (und umgekehrt).¹¹³

Das zweite Prinzip, das Davidson beschreibt, sucht eine generelle Antwort auf die Frage, wann es legitim ist eine Beziehung zwischen zwei Ereignissen als Kausalrelation aufzufassen. Dabei identifiziert er die Gesetzesartigkeit als Kernelement von Verursachungsbeziehungen.

Prinzip des nomologischen Charakters von Kausalität (PNCK): Ereignisse, die in einem Ursache-Wirkungs-Verhältnis zueinander stehen, fallen unter ein striktes deterministisches Gesetz.

Kausalbeziehungen sind laut Davidson also – im Gegensatz etwa zu gesetzesartigen Verallgemeinerungen – dadurch gekennzeichnet, dass die Relation zwischen dem Ursache- und dem Wirkungsereignis stets unter ein striktes deterministisches Gesetz fällt. Natürlich muss das strikte Gesetz, unter das das Ursache- und das Wirkungsereignis fallen, nicht zwangsläufig bekannt sein. Davidson zufolge besteht nämlich ein entscheidender Unterschied „(...) between knowing there is a law ‘covering’ two events and knowing what the law is: (...) singular causal statements entail no law; (...) they entail that there is a law.“¹¹⁴ Die Kenntnis *welches* konkrete strikte Gesetz von einer bestimmten Kausalrelation instanziiert wird, ist nach Davidson zunächst also eher zweitrangig – vielmehr zählt das Wissen, dass *überhaupt* ein

¹¹³ vgl. Davidson, Donald: ‚*Mental Events*‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 208.

¹¹⁴ Davidson, Donald: ‚*Mental Events*‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 160.

solches Gesetz „am Werk“ ist. Wenn nämlich von einer bestimmten Relation bekannt ist, dass sie ein striktes Gesetz instanziiert, dann muss dies zwangsläufig eine *Kausalrelation* sein.

Das dritte Prinzip konstatiert, dass es keine strikten psychophysischen Gesetze gibt, die eine Vorhersage oder Erklärung von mentalen Ereignissen ermöglichen würde.

Prinzip der Anomalie des Mentalen (PAM): Es gibt keine strikten Gesetze anhand derer man mentale Ereignisse erklären oder vorhersagen könnte.

Insgesamt weist Davidson PKI und PNCK zunächst nur einen Status als Annahmen zu, wohingegen er PAM als Tatsache auffasst. Das Problem um die Anomalie des Mentalen kommt nun dadurch zustande, dass einerseits jedes dieser drei Prinzipien für sich genommen zunächst evident erscheint, sie aber andererseits den Eindruck erwecken als wären sie keinesfalls miteinander vereinbar bzw. zusammen kohärent denkbar. Es mutet nämlich an als habe die Wahrheit von zwei beliebigen Prinzipien unweigerlich die Falschheit des jeweils dritten Prinzips zur Folge. Hierzu eine kurze Analyse der möglichen Kombinationen: Sind PKI und PNCK wahr, dann scheint es als müsse PAM falsch sein: denn falls mentale Ereignisse kausal mit physikalischen Ereignissen interagieren und Kausalbeziehungen immer ein striktes deterministisches Gesetz instanziierten, dann hat es den Anschein als müsse es auch strikte Gesetze geben, anhand derer man mentale Ereignisse vorhersagen bzw. erklären kann – was wiederum einer Negation von PAM entspräche. Wenn hingegen PKI und PAM wahr sind, dann erweckt dies den Eindruck, als müsse PNCK zwangsläufig falsch sein. Treten nämlich mentale Ereignisse einerseits mit physikalischen Ereignissen in kausale Wechselwirkung und gilt andererseits zugleich die Anomalie des Mentalen, dann existieren somit Kausalbeziehungen, die nicht unter ein striktes, deterministisches Gesetz fallen. Entsprechen schließlich PNCK und PAM der Wahrheit, so erweckt dies den Eindruck als müsse PKI falsch sein. Denn falls bei Kausalrelationen immer ein striktes, deterministisches Gesetz am Werk ist und strikte psychophysische Gesetze grundsätzlich nicht möglich sind, dann können mentale Ereignisse allem Anschein nach nicht in Kausalbeziehungen mit physikalischen Ereignissen treten.

Davidsons Anomaler Monismus als Lösung für das Problem um die Anomalie des Mentalen

Da Davidson an der Wahrheit aller drei Prinzipien festhalten will, sieht er seine Aufgabe darin den Widerspruch, den sie auf den ersten Blick implizieren, wegzu erklären. Zu diesem Zweck gliedert er den Rest seines Aufsatzes in drei Teile: im ersten versucht Davidson eine

Identitätstheorie von Mentalem und Physikalischen zu entwerfen, die mit der Wahrheit der drei Prinzipien in Einklang gebracht werden kann. Danach argumentiert er, dass es keine strikten psychophysischen Gesetze geben kann, woraus sich mit nicht allzu kritischen Zusatzannahmen das Prinzip der Anomalie des Mentalen (PAM) ableiten lasse. Im letzten Schritt seines Essays versucht Davidson zu zeigen, dass aus dem Fehlen strikter psychophysischer Gesetze und den anderen beiden Prämissen die Wahrheit der von ihm präsentierten Identitätstheorie gefolgert werden kann.

An den Anfang seiner Theorie stellt Davidson sein zentrales Identitätspostulat, sowie seine Definitionen des Ereignisbegriffs und der Identitätsbedingungen für Ereignisse:

(...) mental events are identical with physical events. Events are taken to be unrepeatable, dated individuals such as the particular eruption of a volcano, the (first) birth or death of a person, the playing of the 1968 World Series, or the historic utterance of the words, 'You may fire when ready, Gridley.' We can easily frame identity statements about individual events; examples (true or false) might be:

The death of Scott = the death of the author of *Waverley*;
The assassination of the Archduke Ferdinand = the event that started the First World War;
The eruption of Vesuvius in A.D. 79 = the cause of the destruction of Pompeii.¹¹⁵

Davidsons Kernthese besteht also darin zu behaupten, dass ein beliebiges mentales Ereignis immer mit einem physikalischen Ereignis identisch sein muss. Wobei er unter ‚Ereignis‘ ein nicht wiederholbares, historisch bzw. zeitlich eindeutig datiertes Einzelding versteht, wie etwa einen bestimmten Vulkanausbruch, die Geburt oder den Tod einer Person, das Ausspielen der World Series (der Profi-Baseball-Meisterschaft in den USA) im Jahre 1968 oder die historische Äußerung der Worte ‚Sie können feuern, wenn Sie soweit sind, Gridley‘. Entsprechend einfach lassen sich auch Identitätsaussagen hinsichtlich einzelner Ereignisse formulieren. Als Beispiele können laut Davidson etwa folgende Ereignisse firmieren (wobei der jeweilige Wahrheitswert der entsprechenden Aussagen Davidson zufolge letztlich keine Rolle spielt): Das Ereignis ‚der Tod von Scott‘ ist identisch mit dem Ereignis ‚der Tod vom Autor von *Waverley*‘. Analog ist das Ereignis ‚die Ermordung des Erzherzogs Ferdinand‘ identisch mit dem ‚Ereignis, das den Ersten Weltkrieg ausgelöst hat‘. Schließlich besteht auch eine Identitätsbeziehung zwischen dem Ereignis ‚Ausbruch des Vesuv im Jahr 79 v.Chr.‘ und dem Ereignis ‚die Ursache der Zerstörung Pompeis‘. Alles in allem sieht Davidson in der Identität zweier Einzelereignisse also im Grunde nichts anderes als zwei unterschiedliche Beschreibungen ein und desselben Ereignisses. Dementsprechend sind Ereignisse auch *qua Beschreibung* mental bzw. physikalisch, so Davidson. In diesem Sinne muss auch jedes mentale Ereignis mit einem

¹¹⁵ Davidson, Donald: ‚*Mental Events*‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 210f.

physikalischen Ereignis identisch sein, d.h. jedes Ereignis, das eine mentale Beschreibung besitzt, hat notwendigerweise zugleich auch eine physikalische Beschreibung. Doch was heißt es für ein Ereignis eine mentale bzw. eine physikalische Beschreibung zu besitzen? Dazu Davidsons Definition: „Let us call a description of the form ‘the event that is *M*’ or an open sentence of the form ‘event *x* is *M*’ a *mental description* or a *mental open sentence* if and only if the expression that replaces ‘*M*’ contains at least one mental verb essentially.“¹¹⁶ Ein Ereignis ist folglich nur dann ein mentales Ereignis, wenn es eine mental(istisch)e Beschreibung hat. Ganz analog ist ein physikalisches Ereignis, ein Ereignis, dessen Beschreibung oder dessen offene Sätze im Wesentlichen nur physikal(ist)isches Vokabular heranziehen.

Um eine bessere Einordnung für die von ihm vorgestellte Theorie zu ermöglichen, stellt Davidson folgendes Schema vor, das erläutert wie die Beziehung zwischen mentalen und physikalischen Ereignissen grundsätzlich gedacht werden kann:

It may make the situation clearer to give a fourfold classification of theories of the relation between mental and physical events that emphasizes the independence of claims about laws and claims of identity. On the one hand there are those who assert, and those who deny, the existence of psychophysical laws; on the other hand there are those who say mental events are identical with physical and those who deny this. Theories are thus divided into four sorts: *nomological monism*, which affirms that there are correlating laws and that the events correlated are one (materialists belong in this category); *nomological dualism*, which comprises various forms of parallelism, interactionism, and epiphenomenalism; *anomalous dualism*, which combines ontological dualism with the general failure of laws correlating the mental and the physical (Cartesianism). And finally there is *anomalous monism*, which classifies the position I wish to occupy.¹¹⁷

Die beiden Dimensionen, anhand derer Davidson seine Theorie charakterisiert, sind zum Einen die Frage nach der Existenz psychophysischer Gesetze und zum Anderen die Frage nach der Identität von mentalen und physikalischen Ereignissen. Für Davidson ist es essentiell diese beiden Fragestellungen als vollkommen unabhängig von einander zu betrachten. Ergo sind grundsätzlich vier Kombinationen dieser beiden Merkmale denkbar: Der Nomologische Monismus behauptet, dass es psychophysische Gesetze gibt und dass mentale und physikalische Ereignisse identisch sind. Auch der Nomologische Dualismus geht von der Existenz psychophysischer Gesetzmäßigkeiten aus, bestreitet aber die Identität von mentalen und physikalischen Ereignissen. Der Anomale Dualismus hingegen behauptet eine dualistische Ontologie, leugnet allerdings das Vorhandensein strikter psychophysischer Gesetze. Davidson vertritt schließlich die Position des Anomalen Monismus, die im Wesentlichen besagt, dass es einerseits keine psychophysischen Gesetze gibt, andererseits aber mentale und physikalische Ereignisse identisch sind. Davidsons Anomaler Monismus gleicht dem Materialismus also inso-

¹¹⁶ Davidson, Donald: ‚*Mental Events*‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 211.

¹¹⁷ Davidson, Donald: ‚*Mental Events*‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 213f.

fern er behauptet alle Ereignisse seien im Grunde physikalisch. Allerdings unterscheidet sich der Anomale Monismus von Materialismus darin, dass er die These zurückweist für alle mentalen Phänomene gäbe es eine rein physikalische Erklärung. Somit kann der Anomale Monismus auch kaum als reduktionistische Position gelten, so Davidson.

Trotz des Fehlens psychophysischer Gesetze und der Unmöglichkeit einer vollständigen Reduktion des Mentalen auf das Physikalische, betont Davidson, dass es seiner Theorie gemäß dennoch möglich sei, dass mentale Charakteristika in gewisser Weise von physikalischen Charakteristika abhängig sind bzw. auf ihnen supervenieren. Davidson führt dies weiter so aus: „Such supervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respect, or that an object cannot alter in some mental respect without altering in some physical respect. Dependence or supervenience of this kind does not entail reducibility through law or definition: if it did, we could reduce moral properties to descriptive, and there is good reason to *believe* cannot be done; and we might be able to reduce truth in a formal system to syntactical properties, and this we *know* cannot in general be done.“¹¹⁸ Mentale Ereignisse supervenieren also in dem Sinne auf physikalischen Ereignissen, dass zwei in physikalischer Hinsicht identische Ereignisse sich nicht in einem mentalen Aspekt unterscheiden können. Wenn das Mentale auf diese Art und Weise auf dem Physikalischen superveniert, dann ist es laut Davidson also nicht möglich, dass sich zwei Ereignisse in physikalischer Hinsicht vollständig gleichen, aber in mentaler Hinsicht einen Unterschied aufweisen. Umgekehrt gilt auch, dass eine Änderung auf der mentalen Ebene eine Veränderung im physikalischen Bereich voraussetzt. Diese Form der Abhängigkeit oder Supervenienz könne jedoch nicht mit der Reduzierbarkeit durch ein Gesetz oder durch Definition gleichgesetzt werden, so Davidson. Wenn dem so wäre, müssten moralische Eigenschaften auf deskriptive Eigenschaften reduzierbar sein und es gibt gute Gründe anzunehmen, dass dies keine gangbare Option ist. Zugleich müsste dann auch die Wahrheit in einem formalen System auf dessen syntaktische Eigenschaften reduzierbar sein – hiervon weiß man jedoch, dass dies generell unmöglich ist, so Davidson.

Nachdem Davidson seine Theorie vorgestellt hat, untersucht er deren Verhältnis zu den Prinzipien PKI, PNCK und PAM und gelangt diesbezüglich zu folgendem Schluss:

It should now be evident how anomalous monism reconciles the three original principles. Causality and identity are relations between individual events no matter how described. But laws are linguistic; and so events can instantiate laws, and hence be explained or predicted in the light of laws, only as those events are described in one or another way. The principle of causal interaction deals with events in extension and is therefore blind to the mental-physical dichotomy. The principle of the anomalism of the mental concerns events described as mental,

¹¹⁸ Davidson, Donald: ‚*Mental Events*‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 214.

for events are mental only as described. The principle of the nomological character of causality must be read carefully: it says that when events are related as cause and effect, they have descriptions that instantiate a law. It does not say that every true singular statement of causality instantiates a law.¹¹⁹

Der Anomale Monismus ermögliche also ein kohärentes Vertreten der drei anfangs diskutierten Prinzipien, so Davidson. Denn sowohl unter ‚Kausalität‘ und als auch unter ‚Identität‘ ist jeweils eine Beziehung zwischen Einzelereignissen – und zwar unabhängig von deren Art der Beschreibung – zu verstehen. Gesetze hingegen sind sprachlicher Natur. Aus diesem Grund können Ereignisse nur dann Gesetze instanziierten und dementsprechend auch nur dann mittels dieser Gesetze erklärt bzw. vorhergesagt werden, wenn sie auf eine bestimmte Art und Weise beschrieben werden. Das Prinzip der kausalen Interaktion bezieht sich unmittelbar auf die Extension von Ereignissen. Daraus folgt, dass die Art der Beschreibung eines Ereignisses – also ob es als ‚mental‘ oder als ‚physikalisch‘ etikettiert wird – keine Rolle für PKI spielt. Das Prinzip der Anomalie des Mentalen hingegen, referiert naturgemäß ausschließlich auf mentale Ereignisse, d.h. PAM ist direkt abhängig von der Art der Beschreibung eines Ereignisses – schließlich sind mentale Ereignisse Davidson zufolge nur deshalb mental, weil sie eine mentale Beschreibung besitzen. Das Prinzip des nomologischen Charakters von Kausalität wiederum behauptet, dass es für zwei Ereignisse, die in einem Verhältnis von Ursache und Wirkung zueinander stehen, eine Beschreibung gibt, die ein Gesetz instanziiert. Allerdings besagt PNCK nicht, dass jede wahre Einzelaussage über Kausalität ein solches Gesetz instanziierten muss.

So kann es durchaus der Fall sein, dass ein Ereignis, das mental beschrieben wird, ein anderes Ereignis, das eine physikalische Beschreibung besitzt, verursacht und damit PKI erfüllt ist. Ferner ist es nach Davidson auch möglich, dass es kein striktes Gesetz gibt, mithilfe dessen man die Kausalrelation zwischen dem mentalen Ursacheereignis und dem physikalischen Wirkungereignis erklären könnte oder hätte voraussagen können – somit gilt auch PAM. Entgegen der ursprünglichen Intuition muss dies laut Davidson jedoch nicht die Falschheit von PNCK zur Folge haben. Denn für das in Frage stehende Ursache- und Wirkungereignis kann es zweifellos ein striktes Gesetz geben, das bei dieser Kausalrelation instanziiert wird – allerdings kann bei Ausformulierung dieses Gesetzes nicht die anfängliche Beschreibung der beiden Ereignisse beibehalten werden. Erhält das zunächst mental(istisch) beschriebene Ursachenereignis nämlich eine physikalis(tis)che¹²⁰ Beschreibung, so ermöglicht dies das Ausfor-

¹¹⁹ Davidson, Donald: ‚*Mental Events*‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 215.

¹²⁰ Aufgrund von Davidsons zentralem Identitätspostulat muss jedes mentale Einzelereignis mit einem physikalischen Ereignis identisch sein. Da Ereignisse nur dank ihrer jeweiligen Beschreibung mental oder physikalisch sind, muss es folglich für jedes mentale Ereignis auch eine physikalische Beschreibung geben.

mulieren eines strikten Gesetzes für die in Frage stehende Kausalrelation. Infolgedessen wird dergestalt auch PNCK entsprochen. Durch diese Argumentationslinie erachtet Davidson auch den zuerst diagnostizierten Widerspruch bezüglich der konjunkten Annahme von PKI, PNCK und PAM als obsolet. Allen anfänglichen Bedenken zuwider, sei es anhand seiner Theorie also doch möglich die drei Prinzipien zusammen auf eine stimmige Art und Weise zu vertreten, so Davidson. An einem Beispiel verdeutlicht könnte Davidsons Gedankengang in etwa wie folgt rekonstruiert werden: Angenommen ein mentales Ereignis, z.B. Freds Wunsch nach einem kühlen Bier, verursacht ein physikalisches Ereignis, etwa, dass Fred zum Kühlschrank läuft. Da in diesem Fall ein mentales Ereignis ein physikalisches Ereignis verursacht, gilt PKI. Ein striktes deterministisches Gesetz, das eine Erklärung oder Vorhersage von Freds Gang zum Kühlschrank aufgrund seines Wunsches nach einem kühlen Bier ermöglicht, ist jedoch kaum vorstellbar. So könnte Fred beispielsweise den Wunsch nach einem kühlen Bier verspüren, aber gleichzeitig auch wissen, dass in seinem Kühlschrank zurzeit kein Bier ist. Oder Fred könnte sich zu dem Zeitpunkt, an dem er das Verlangen nach einem Bier verspürt, in einem anderen Teil seines Hauses befinden – dann wäre das physikalische Wirkungsereignis ein anderes (auch wenn es tatsächlich ein Zum-Kühlschrank-Gehen wäre). Eine weitere vorstellbare Alternative wäre, wenn Fred zwar ein kühles Bier begehrt, andererseits aber den Vorsatz gefasst hat kein Bier mehr zu trinken. Tatsächlich kann man eine Vielzahl solcher Szenarien konstruieren, was die Hoffnung auf das Finden eines strikten Gesetzes, das bei der besagten Kausalbeziehung zwischen Freds Wunsch nach einem kühlen Bier und seinem Gang zum Kühlschrank, instanziiert wird, begräbt. Insofern ist auch PAM erfüllt. Nimmt man jedoch an, dass Freds Wunsch nach einem kühlen Bier nur qua Beschreibung ein mentales Ereignis ist und dass ein extensionsgleiches physikalisches Prädikat zur Beschreibung dieses Ereignisses existiert, dann könnte es mithilfe dieser physikalischen Beschreibung des Ereignisses möglich sein, ein striktes Gesetz abzuleiten, das durch diese Kausalrelation instanziiert wird. In diesem Fall wäre auch PCNK bestätigt.

Der zweite von Davidson in *Mental Events* diskutierte Themenkomplex bezieht sich auf die Nicht-Existenz psychophysischer Gesetze. In diesem Zusammenhang stellt er zunächst sein Verständnis von der Nicht-Reduzierbarkeit des Mentalen vor: „The thesis is (...) that the mental is nomologically irreducible: there may be *true* general statements relating the mental and the physical, statements that have the logical form of a law; but they are not *lawlike* (...).“¹²¹ Davidson betrachtet das Mentale also als nomologisch nicht reduzierbar. D.h. es kann zwar durchaus wahre Allgemeinaussagen geben, die Mentales und Physikalisches zu-

¹²¹ Davidson, Donald: ‚*Mental Events*‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 216.

einander in Bezug setzen, einen gesetzesartigen Charakter können diese Generalisierungen jedoch nie besitzen. Die Definition für gesetzesartige Aussagen lautet hierbei wie folgt: „Lawlike statements are general statements that support counterfactual and subjunctive claims, and are supported by their instances.“¹²² Dementsprechend sind gesetzesartige Aussagen Allgemeinaussagen, die durch kontrafaktische und konjunktivistische Behauptungen gestützt und von ihren Instanzen bestätigt werden. Einen ersten Hinweis darauf, warum solche gesetzesartigen Aussagen über Relationen zwischen Mentalem und Physikalischem zumindest als sehr unplausibel erscheinen, sieht Davidson im Holismus des Mentalen, insbesondere im Hinblick auf Verhalten, begründet: „Beliefs and desires issue in behaviour only as modified and mediated by further beliefs and desires, attitudes and attendings, without limit. Clearly this holism of the mental realm is a clue to both to the autonomy and to the anomalous character of the mental.“¹²³ Im Allgemeinen sind Annahmen und Wünsche, die ein bestimmtes Verhalten verursachen, in ein fein verästeltes Gefüge von weiteren Glaubenssätzen, Anliegen, Einstellungen, Erwartungen, usw., die einander wechselseitig beeinflussen, eingebettet. Dieses Beziehungsgeflecht erweckt eine tiefschürfende Skepsis gegenüber der Annahme von wahren, präzisen Allgemeinaussagen, die Gedanken und Verhalten zueinander in Relation setzen könnten. Wenn ein mentales Ereignis einer bestimmten (mentalen) Art für gewöhnlich mit einem Ereignis einer gewissen physikalischen Art einhergeht, so ist dies zwar gewiss oftmals ein guter Grund dafür anzunehmen, dass ähnlich gelagerte Fälle auch diesem Beispiel folgen werden, jedoch sind Verallgemeinerungen und praktisches Wissen dieses Typs immer nur annäherungsweise wahr. Ihre Hauptaufgabe besteht vornehmlich darin, singulären Kausalaussagen Nachdruck zu verleihen und für die darunter subsummierbaren Einzelereignisse eine Erklärung liefern zu können, so Davidson. In unserem ‚Alltagsgeschäft‘ hinsichtlich der Erklärung und Vorhersage von Ereignissen bedienen wir uns ständig solcher ungenauen Generalisierungen, entweder weil die exakten Gesetze nicht bekannt sind oder weil uns die Beschreibungen der Ereignisse, die die Relevanz eines solchen Gesetzes nachweisen könnten, fehlen. Allerdings gelte es grundsätzlich zwei verschiedene Arten von solchen Verallgemeinerungen zu unterscheiden:

In our daily traffic with events and actions that must be foreseen and understood, we perforce make use of sketchy summary generalization, for we do not know a more accurate law, or if we do, we lack a description of the particular events in which we are interested that would show the relevance of the law. But there is an important distinction to be made within the category of the rule of thumb. On the one hand, there are generalizations whose positive instances give us reason to believe the generalization itself could be improved by adding

¹²² Davidson, Donald: ‚*Mental Events*‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 217.

¹²³ Davidson, Donald: ‚*Mental Events*‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 217.

further provisios and conditions stated in the same general vocabulary as the original generalization. Such a generalization points to the form of vocabulary of the finished law: we may say that it is a *homonomic* generalization. On the other hand there are generalizations which when instantiated may give us reason to believe there is a precise law at work, but one that can be stated only by shifting to a different vocabulary. We may call such generalizations heteronomic. I suppose most of our practical lore (and science) is *heteronomic*.¹²⁴

Zum Einen gibt es nach Davidson also homonomische Generalisierungen, die dadurch gekennzeichnet sind, dass sie durch zusätzliche Klauseln und Bedingungen sozusagen zu einem regelrechten Gesetz ‚ausgebaut‘ bzw. erweitert werden können. In anderen Worten bedient sich bereits die ursprüngliche Formulierung der Generalisierung desselben Wortschatzes wie später das endgültige Gesetz. Das Gegenstück zu homonomischen Generalisierungen nennt Davidson heteronomische Verallgemeinerungen. Bei dieser Art von Allaussage muss bei der Ausarbeitung des jeweiligen Gesetzes im Gegensatz zu homonomischen Generalisierungen auf ein grundsätzlich anderes Vokabular als in der anfänglichen Verallgemeinerung zurückgegriffen werden. Die allermeisten Generalisierungen, die in Wissenschaft und Alltag Verwendung finden, gehören Davidson zufolge der Klasse der heteronomischen Allaussagen an. Dies ist wiederum darauf zurückzuführen, dass homonomische Verallgemeinerungen ihre Begrifflichkeiten ausschließlich aus einer umfassenden, in sich geschlossenen Theorie rekrutieren können, wobei diese ultimative Theorie letztendlich entweder deterministisch oder probabilistisch sein kann. Allerdings sei die Physik die einzige Wissenschaft, die den Ansprüchen hinsichtlich homonomischer Generalisierungen (auch nur annähernd) gerecht werden kann, so Davidson. Dementsprechend können Allgemeinaussagen, die sowohl physikalisches als auch mentalistisches Vokabular enthalten nur heteronomisch sein. Insofern erkläre sich daraus auch warum es keine strikten psychophysischen Gesetze geben kann. Davidson fasst seinen Gedankengang so zusammen:

The heteronomic character of general statements linking the mental and the physical traces back to this central role of translation in the description of all propositional attitudes, and to the indeterminacy of translation. There are no strict psychophysical laws because of the disparate commitments of the mental and physical schemes. It is a feature of physical reality that physical change can be explained by laws that connect it with other changes and conditions physically described. It is a feature of the mental that the attribution of mental phenomena must be responsible to the background of reasons and beliefs, and intentions of the individual. There cannot be tight connections between the realms if each is to retain allegiance to its proper source of evidence. The nomological irreducibility of the mental does not derive merely from the seamless nature of the world of thought, preference, and intention, for such interdependence is common to physical theory, and is compatible with there being a single right way of interpreting a man's attitudes without relativization to a scheme of translation. Nor is the irreducibility due simply to the possibility of many equally eligible schemes, for it is compatible with an arbitrary choice of one scheme relative to which assignments of mental traits are made. The point is rather that when we use the concepts of belief, desire, and the rest, we must stand prepared, as the evidence accumulates, to adjust our theory in the light of considerations of overall cogency: the constitutive ideal of rationality partly controls each phrase in the evolution of what must be an evolving theory. An arbitrary choice of translation scheme would preclude such opportunistic tempering of theory; put differently, a right arbitrary choice of translation

¹²⁴ Davidson, Donald: ‚*Mental Events*‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 219.

manual would be of a manual acceptable in the light of all possible evidence, and this is a choice we cannot make. We must conclude, I think, that nomological slack between the mental and the physical is essential as long as we conceive of man as a rational animal.¹²⁵

Da das mentale und das physikalische Erklärungsmodell jeweils komplementäre und somit nicht vereinbare Foki setzen, kann es keine strikten psychophysischen Gesetze geben, so Davidson. Während es ein Merkmal der physikalischen Wirklichkeit(skonstruktion) ist, dass sich physikalische Veränderungen durch Gesetze erklären lassen, die diese dynamischen Prozesse in Bezug zu anderen physikalisch beschreibbaren Veränderungen oder Bedingungen setzen, gilt dies umgekehrt auch genauso sehr für das Mentale. So können mentale Phänomene nur vor dem Hintergrund von Gründen, Glaubenssätzen und Intentionen eines Handelnden verstanden werden. Da beide Bereiche allein ihre jeweils ureigenen Quellen als beweiskräftig ansehen, kann es keine enge Verbindung zwischen der mentalen und der physikalischen Beschreibungsebene geben. Die nomologische Nicht-Reduzierbarkeit des Mentalen rührt deswegen nicht allein daher, dass die Welt des Mentalen in sich abgeschlossen ist. Denn eine solche strenge Interdependenz ist auch aus der Physik bekannt und wäre durchaus mit der These vereinbar, dass es nur eine einzige richtige Methode zur Deutung der Einstellungen eines Menschen gibt, die ohne jeglichen Bezug auf ein Übersetzungsmodell gültig wäre. Tatsächlich ist eine korrekte Interpretation der Haltungen einer Person laut Davidson aber nur aus der mentalen Perspektive möglich. Die bloße In-Sich-Abgeschlossenheit der mentalen Welt habe also nicht deren explanatorische Autarkie zur Folge und beweise deswegen auch nicht die nomologische Nicht-Reduzierbarkeit des Mentalen. Genauso wenig könne man die nomologische Nicht-Reduzierbarkeit des Mentalen von der bloßen Möglichkeit mehrerer gleichermaßen gültiger Übersetzungsmodelle herleiten. Schließlich wäre ansonsten eine willkürliche Zuordnung irgendeines Übersetzungsschematas für das Zuschreiben mentaler Charakteristika möglich. Eine solche Attribution macht Davidson zufolge allerdings nur in einem mental beschriebenen Kontext Sinn. Der eigentliche Grund für die nomologische Nicht-Reduzierbarkeit des Mentalen liege vielmehr im Verständnis des Menschen als ein rationales Wesen. Gebraucht man nämlich mentale Begriffe wie Glauben, Verlangen, etc. in einem Erklärungsmodell, dann sollte man auch bereit sein im Lichte neuer Erkenntnisse die Theorie gegebenenfalls an veränderte Gegebenheiten anzupassen, um eine klare Stringenz der Theorie aufrechtzuerhalten. Dies wiederum erweist sich als höchst unverträglich mit der beliebigen Zuweisung eines Übersetzungsmodells, weil darunter die Adaptierbarkeit der Theorie an veränderte Umstände sehr litte. Daraus zieht Davidson den Schluss, dass der nomologische Spalt

¹²⁵ Davidson, Donald: 'Mental Events' in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 222f.

zwischen Mentalem und Physikalischem solange Bestand hat, wie der Mensch als eine rationale Kreatur verstanden wird. Folglich bestehe ein kategorialer Unterschied zwischen dem Mentalen und dem Physikalischen.

Im letzten Abschnitt seines Essays versucht Davidson nachzuweisen, dass das Fehlen psychophysischer Gesetze und die Akzeptanz von PKI und PAM notwendigerweise die Wahrheit seiner Theorie zur Folge hat. Die Besonderheit seines Ansatzes sieht er darin herauszustellen, dass die Identität zwischen Mentalem und Physikalischem gerade der Gesetzlosigkeit des Mentalen zu verdanken ist. Dies ist insofern überraschend, als dass das Physikalische im Allgemeinen als das Exempel des Gesetzmäßigen schlechthin gilt. Der Gedankengang, der ihn zu dieser Schlussfolgerung führt, nimmt seinen Ausgang vom Prinzip der kausalen Abhängigkeit des Mentalen (oder von PKI), demzufolge zumindest einige mentale Ereignisse die Ursachen oder Wirkungen von physikalischen Ereignissen sind. Nach PNCK gibt es für jede wahre, singuläre Kausalaussage ein striktes Gesetz, das Ursache- und Wirkungsereignis miteinander verbindet. Eine homonomische Generalisierung – sozusagen die Vorstufe eines Gesetzes – muss ihre Begriffe aus einem umfassenden, geschlossenen System beziehen. Solche Verallgemeinerungen wiederum sind im Grunde nur in der Physik vorstellbar. Schließlich scheint das Mentale nicht in derselben Art und Weise den Anspruch auf theoretischer In-Sich-Geschlossenheit erheben zu können wie das Physikalische. Denn es gibt nämlich zahlreiche äußere Einflussfaktoren, die selbst kein systematischer Bestandteil des Mentalen sind, es aber trotzdem kausal beeinflussen. Treffen diese Beobachtungen zu, dann ist es prinzipiell nicht möglich, dass psychophysische Aussagen einen integralen Bestandteil eines strikten Gesetzes bilden können. Daraus lässt sich wiederum ableiten, dass es kein striktes Gesetz gibt anhand dessen man mentale Phänomene erklären oder vorhersagen könnte. Diese Aussage ist jedoch nichts anderes als PAM. Hieraus lässt sich Davidson zufolge die Identität von mentalen Ereignissen mit physikalischen Ereignissen mühelos deduzieren:

The demonstration of identity follows easily. Suppose *m*, a mental event, caused *p*, a physical event; then, under some description *m* and *p* instantiate a strict law. This law can only be physical (...). But if *m* falls under a physical law, it has a physical description; which is to say it is a physical event. An analogous argument works when a physical event causes a mental event. So every mental event that is causally related to a physical event is a physical event. In order to establish anomalous monism in full generality it would be sufficient to show that every mental event is cause or effect of some physical event; I shall not attempt this.¹²⁶

Nach Davidsons Argumentation muss also gelten, dass wenn ein mentales Ereignis *m* ein physikalisches Ereignis *p* verursacht, es eine Beschreibung von *m* und *p* gibt, die ein striktes Gesetz instanziiert. Da ausschließlich homonomische Generalisierungen zu strikten Gesetzen

¹²⁶ Davidson, Donald: ‚*Mental Events*‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 224.

weiterentwickelt werden können und die Physik der einzige Gegenstandsbereich ist, in dem die Hoffnung besteht diese Art von Verallgemeinerungen zu finden, muss das strikte Gesetz, unter das *m* und *p* fallen, ein physikalisches Gesetz sein. Wenn nun aber *m* unter ein physikalisches Gesetz fällt, so benötigt *m* auch eine physikalische Beschreibung. Schließlich ist die Art der Beschreibung von Ereignissen für das Fallen unter bestimmte Gesetze wesentlich. Besitzt *m* aber auch eine physikalische Beschreibung, so folgt daraus unmittelbar, dass *m* zugleich ein physikalisches Ereignis sein muss, weil Ereignisse nur dank ihrer Beschreibung mental oder physikalisch sind. Ergo muss das mentale Ereignis *m* mit einem (beliebigen) physikalischen Ereignis (etwa *p**) identisch sein. Da ein analoges Argument auch für den umgekehrten Fall – i.e. wenn ein physikalisches Ereignis ein mentales Ereignis verursacht – gültig ist, lässt sich dieses Prinzip auch sehr einfach verallgemeinern. Dementsprechend muss jedes mentale Ereignis, das mit einem physikalischen Ereignis kausal interagiert eine physikalische Beschreibung besitzen und daher mit einem physikalischen Ereignis identisch sein. Um Allgemeingültigkeit zu erlangen, müsste ein Advokat des Anomalen Monismus dementsprechend lediglich nachweisen, dass jedes mentale Ereignis die Ursache oder Wirkung eines physikalischen Ereignisses ist, so Davidson. Diesen Anspruch erhebt er jedoch nicht – ihm genüge es die Gültigkeit des Anomalen Monismus auf diejenigen Fälle zu beschränken, in denen mentale Ereignisse tatsächlich mit physikalischen Ereignissen in kausale Wechselwirkung treten.

Immer wenn ein Ereignis ein anderes Ereignis verursacht, instanziiieren diese beiden Ereignisse ein striktes Gesetz, genauer gesagt gilt dies nur falls sie auf adäquate Weise beschrieben werden. In der Praxis ist in den meisten Fällen entweder das strikte Gesetz oder die relevante Art der Beschreibung der Ereignisse jedoch unbekannt. Trotzdem kann man Wissen über solche singulären Kausalrelationen besitzen. Dieses Wissen ist in Form von ungefähren, heteronomischen Verallgemeinerungen begründet. Solche Generalisierungen sind nämlich insofern gesetzesartig als ihr Eintreten es vernünftig erscheinen lässt, dass künftige Instanzen diesem Beispiel folgen werden. Sie sind jedoch nicht in dem Sinne gesetzesartig, dass sie beliebig verfeinert werden könnten. Übertragen auf Fälle mentaler Verursachung bedeutet das: „(...) we see that it is possible to know that a mental event is identical with some physical event without knowing which one (in the sense of being able to give a unique physical description that brings it under a relevant law). Even if someone knew the entire physical history of the world, and every mental event were identical with a physical, it would not follow that he

could predict or explain a single mental event (so described, of course).“¹²⁷ Grundsätzlich ist es bei Instanzen mentaler Verursachung also durchaus möglich zu wissen, dass das in Frage stehende mentale Ursache- oder Wirkungsereignis identisch mit einem physikalischen Ereignis ist, ohne dabei angeben zu können mit welchem physikalischen Ereignis es genau identisch ist. Mit anderen Worten muss man im konkreten Fall nicht in der Lage sein können die explizite physikalische Beschreibung des mentalen Ereignisses zu verlautbaren, die es ermöglicht diese singuläre Kausalrelation unter ein striktes Gesetz zu subsumieren. Angenommen, jemand würde die gesamte physikalische Geschichte der Welt kennen und jedes mentale Ereignis wäre identisch mit einem physikalischen Ereignis, so würde daraus aber keineswegs folgen, dass diese Person auch nur ein einziges mentales Ereignis vorhersagen oder erklären könnte. Natürlich könnte ein solcher, im Bereich der Physik, Allwissender das betreffende Ereignis *als physikalisches* Ereignis voraussagen, d.h. unter seiner physikalischen Beschreibung, es wäre ihm jedoch nicht möglich das besagte Ereignis *als mentales* Ereignis vorherzusagen, d.h. die mentale Beschreibung dieses Ereignisses kundzutun. Auf diese Art und Weise könne man schließlich der Autonomie des Mentalen gerecht werden, so Davidson.

Letztendlich resümiert Davidson seine Überlegungen dazu wie der vermeintliche Widerspruch bezüglich mentaler Verursachung aufgelöst werden könne folgendermaßen:

Two features of mental events in their relation to the physical – causal dependence and nomological independence – combine, then, to dissolve what has often seemed a paradox, the efficacy of thought and purpose in the material world, and their freedom from law. When we portray events as perceivings, rememberings, decisions and actions, we necessarily locate them amid physical happenings through the relation of cause and effect; but as long as we do not change the idiom that same mode of portrayal insulates mental events from strict laws that can in principle be called upon to explain and predict physical phenomena.

Mental events as a class cannot be explained by physical science; particular mental events can when we know particular identities. But the explanations of mental events in which we are typically interested relate them to other mental events and conditions. We explain a free man's action, for example, by appeal to his desires, habits, knowledge and perceptions. Such accounts of intentional behaviour operate in a conceptual framework removed from the direct reach of physical law, by describing both cause and effect, reason and action, as aspects of a portrait of a human agent. The anomalism of the mental is thus a necessary condition for viewing action as autonomous.¹²⁸

Was zunächst als Paradox hinsichtlich mentaler Ereignisse erschien – nämlich die kausale Wirksamkeit von Gedanken und Absichten in einer im Grunde rein physikalischen Welt einerseits und deren Unvermögen unter strikte Gesetze zu fallen andererseits – löst sich wegen zwei Merkmalen des Mentalen auf: diese sind zum Einen die kausale Abhängigkeit des Mentalen und zum Anderen seine nomologische Unabhängigkeit. Zumindest in der Alltagssprache werden Ereignisse, die z.B. als Wahrnehmungen, Entscheidungen oder Handlungen – also als

¹²⁷ Davidson, Donald: ‚*Mental Events*‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 224.

¹²⁸ Davidson, Donald: ‚*Mental Events*‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001, S. 224f.

mentale Ereignisse – dargestellt werden, häufig als Ursache oder Wirkung in physikalische Kausalketten eingereiht. Solange man nun aber in der mentalen Redeweise verharret, werden mentale Ereignisse dadurch generell aus strikten Gesetzen ausgeschlossen, die es ermöglichen physikalische Ereignisse vorherzusagen oder zu erklären. Somit stehen mentale Ereignisse einerseits in kausaler Abhängigkeit zu physikalischen Ereignissen, andererseits genießen sie jedoch auch nomologische Unabhängigkeit von diesen. Aufgrund der nomologischen Nicht-Reduzierbarkeit mentaler Ereignisse kann die Klasse dieser Ereignisse auch durch die Physik nicht (vollständig) erklärt werden – zumindest als Klasse. So ist es zwar durchaus möglich einzelne mentale Ereignisse bei Kenntnis des jeweils identischen physikalischen Ereignisses bzw. der physikalischen Beschreibung zu erklären oder vorherzusagen. Allerdings beschränkt sich diese Möglichkeit zur Vorhersage bzw. Erklärung nur auf die *physikalische* Beschreibung des Ereignisses – unter seiner mentalen Beschreibung ist dies hingegen nicht möglich. Die Erklärungen von mentalen Ereignissen, an denen wir in der Regel interessiert sind, setzen diese allerdings in Bezug zu anderen mentalen Ereignissen. Dementsprechend werden etwa freie Handlungen einer Person unter Berufung auf deren Wünsche, Wahrnehmungen, Wissen und Gewohnheiten erklärt. Diese Art Verhalten über die Intentionen eines Handelnden zu explizieren, entzieht sich jedoch kraft ihres begrifflichen Bezugssystems dem direkten Wirkungskreis physikalischer Gesetzmäßigkeiten, indem es Ursache und Wirkung als Gesichtspunkte einer handelnden Person sieht. Schlussendlich ist also in der Anomalie des Mentalen eine notwendige Bedingung dafür zu sehen, Handlungen als autonom anzusehen.

Ist das Mentale qua Mentalem kausal relevant?

In seinem Aufsatz *The Argument for Anomalous Monism* nimmt Ted Honderich kritisch Stellung zu Davidsons Argument für den Anomalen Monismus. Laut Honderich besteht sozusagen der ‚Clou‘ von Davidsons Gedankengang darin, den Widerspruch zwischen PKI, PNCK und PAM aufzuheben, indem er eine besondere Lesart von PNCK anbietet. Nach dieser Interpretation gilt es die kausale Verbindung zwischen zwei Ereignissen einerseits und deren Art der Beschreibung andererseits zu unterscheiden. So besteht die kausale Verbindung zwischen zwei Ereignissen unabhängig von deren Beschreibung. Für die Frage hingegen, ob die beiden Ereignisse in einer gesetzesartigen Verbindung zueinander stehen, spielt die jeweilige Beschreibung eine zentrale Rolle. Folglich können ein mentales Ereignis und ein physikalisches Ereignis de facto durchaus in einem Kausalverhältnis zueinander stehen. Es ist jedoch möglich, dass die beiden Ereignisse (unter dieser Beschreibung) kein striktes Gesetz instanzieren.

Nimmt man zudem an, dass es keine strikten psychophysischen Gesetze geben kann (d.h. falls man von der Wahrheit von PAM ausgeht), dann muss ein mentales Ereignis, das mit einem physikalischen Ereignis in einem Kausalverhältnis steht, mit einem (anderen) physikalischen Ereignis identisch sein. Die Beziehung zwischen dem mentalen und dem physikalischen Ereignis habe daher die Form einer: „(...) lawless identity. It is not a matter of law that the mental event is what it is, identical with a physical event. As the idea of Anomalous Monism is also expressed, on the assumption that types bring in law, we have token-identity but not type-identity.“¹²⁹ Die Identitätsbeziehung zwischen dem mentalen und dem physikalischen Ereignis ist also nicht gesetzesartig in dem Sinn, dass das Wesen des mentalen Ereignisses (nämlich mit einem physikalischen Ereignis identisch zu sein) einem Gesetz zu verdanken wäre. Unter der Annahme, dass für die Formulierung von Gesetzen Typen herangezogen werden, kann der Kerngedanke von Davidsons Anomalen Monismus auch als Token-Identitätstheorie wiedergegeben werden. Denn jedes mentale Ereignis-Token muss mit einem physikalischen Ereignis-Token identisch sein, zwischen mentalen und physikalischen Ereignis-Typen hingegen muss eine solche Identität jedoch nicht bestehen.

Um zu einem tieferen Verständnis von Davidsons Theorie zu gelangen und als Ausgangspunkt für seine Kritik an diesem Ansatz, bedient sich Honderich folgendes Beispiels:

Certain it *is* true that when I put some pears on a scale, something green and French did cause the pointer to move to the two-pound mark, but there in fact is no entailed law connecting greenness and Frenchness with the pointer's so moving. There is in fact no law at all connecting the event in virtue of its being of something green and French with the pointer's moving to the two-pound mark. There is no lawlike connection connecting the first event in virtue of greenness and Frenchness with the second event in virtue of its being the pointer's moving to the two-pound mark.

It is to be noticed that we have given clear sense to talk of something's being such and such *as something or other*, or *under a description*. To talk this way is to speak of certain properties of a thing rather than others. To say two things are not in lawlike connection under certain descriptions is to say that certain of their properties are not in lawlike connection in virtue of certain of their properties.¹³⁰

Sicherlich ist es (unter bestimmten Umständen) wahr, dass wenn man einige Birnen auf eine Waagschale legt, etwas Grünes und Französisches den Ausschlag des Zeigers der Waage bis zur Zwei-Pfund-Marke verursacht. Allerdings zieht dies gewiss kein Gesetz nach sich, das die grüne Farbe und die französische Herkunft der Birnen mit der Bewegung des Zeigers bis zur Zwei-Pfund-Marke in Verbindung bringt. In anderen Worten gibt es keinen gesetzesartigen Zusammenhang zwischen dem ersten Ereignis (i.e. dem Legen von grünen, französischen Birnen auf eine Waage) aufgrund der ‚Grünheit‘ und der ‚Französischheit‘ und dem zweiten Ereignis (i.e. dem Ausschlag des Zeigers der Waage bis zur Zwei-Pfund-Marke) in seiner

¹²⁹ Honderich, Ted: ‚*The Argument for Anomalous Monism*‘ in: *Analysis*, Vol. 42, 1982, S. 60.

¹³⁰ Honderich, Ted: ‚*The Argument for Anomalous Monism*‘ in: *Analysis*, Vol. 42, 1982, S. 60f.

Funktion als Bewegung des Zeigers zur Zwei-Pfund-Marke. Versteht man Davidsons Rede-weise von ‚unter einer bestimmten Beschreibung‘ ferner als ein Sprechen über Eigenschaften, dann bedeutet ‚zwei Dinge stehen unter einer gewissen Beschreibung nicht in einer gesetz-artigen Beziehung zueinander‘, nichts anderes als, dass aufgrund bestimmter Eigenschaften diese Dinge nicht in einem gesetzartigen Zusammenhang stehen. Auf das Birnen-Beispiel bezogen heißt das: „(...) that neither the greenness nor the Frenchness of the pears was rele-vant to the event of the pears’ being put on the scale in so far as that event caused the pointer to move to the two-pound mark.“¹³¹ Kurzum besaßen also weder die grüne Farbe noch der französische Ursprung der Birnen Relevanz für das In die Waageschalelegen der Früchte, insofern als dass dieses Ereignis die Ursache für das Ausschlagen des Zeigers der Waage war. Hinter dieser These steht die Annahme, dass ein Ereignis aufgrund bestimmter Eigenschaften die Ursache ist, die es ist – andere Eigenschaften dieses Ereignisses sind dafür wiederum nicht relevant. Könnte man gegen diese Hypothese ausgehend vom Birnen-Beispiel nicht auch umgekehrt argumentieren? Nämlich in etwa so: das Ereignis des Legens der Birnen auf die Waagschale wäre nicht dasselbe Ereignis gewesen, wenn die Birnen nicht grün und fran-zösisch gewesen wären. Daher kann man auch nicht sagen, dass *dieses* Ereignis nicht die Ur-sache der Bewegung des Zeigers der Waage gewesen wäre. *Dieses* Ereignis wäre nämlich schlichtweg nicht eingetreten, wenn die Birnen nicht grün und französisch gewesen wären. Aus diesem Grund waren die grüne Farbe und der französische Herkunft der Birnen tatsäch-lich kausal relevant für den Ausschlag des Zeigers bis zur Zwei-Pfund-Marke. Dies wäre Honderich zufolge jedoch höchst kontraintuitiv. Denn: „Certain conditional connections hold between the weight of the pears and the pointer’s movement, and they do not hold between the greenness or Frenchness of the pears and the pointer’s movement. The greenness and Frenchness were necessary to the event’s being the event it was, but not necessary to the event’s being the cause it was. Certainly, it may be said that the cause there was would not have existed if the pears have not been green and French. That is consistent with the green-ness and Frenchness being causally irrelevant to the effect.“¹³² Nach Honderich bestehen also bestimmte konditionale Verbindungen zwischen dem Gewicht der Birnen und der Bewegung des Zeigers der Waage, wohingegen kein solcher Zusammenhang zwischen der grünen Farbe oder dem französischen Anbaugebiet der Birnen und dem Ausschlag des Zeigers besteht. Hät-ten die Birnen beispielweise sechs Pfund gewogen, dann hätte sich auch der Zeiger nicht nur bis zur Zwei-Pfund-Marke bewegt. Wenn die Birnen hingegen rot und aus Italien, jedoch gleich schwer, gewesen wären, dann wäre der Zeiger nichtsdestotrotz zur Zwei-Pfund-Marke

¹³¹ Honderich, Ted: ‚*The Argument for Anomalous Monism*‘ in: *Analysis*, Vol. 42, 1982, S. 61.

¹³² Honderich, Ted: ‚*The Argument for Anomalous Monism*‘ in: *Analysis*, Vol. 42, 1982, S. 61f.

gewandert. Deshalb sind die Grünheit und die Französischheit der Birnen sicherlich ein notwendiger Bestandteil dafür, dass dieses Ereignis das *Ereignis* ist, das es ist – allerdings gewiss nicht dafür, dass dieses Ereignis die *Ursache* ist, die es ist. Natürlich könnte man nun behaupten, die Ursache, die eingetreten ist, hätte nicht existiert, wenn die Birnen nicht grün und französisch gewesen wären. Dies ist jedoch konsistent respektive gleichbedeutend damit, die Grünheit und Französischheit der Birnen als kausal irrelevant für die Anzeige der Waage zu erachten. Außerdem ist die Tatsache, dass man sagen würde die Ursache hätte nicht existiert, wenn die Birnen nicht grün und französisch gewesen wären, dem allgemeinen Sprachgebrauch entliehen und nicht irgendeinem Faktum über die kausale Notwendigkeit aller Eigenschaften der Birnen geschuldet – denn ein solches Faktum existiert schlichtweg nicht. Abschließend zieht Honderich folgendes Resümee hinsichtlich der kausalen Relevanz bestimmter Eigenschaften von Ereignissen, die in gesetzesartigen Verbindungen zueinander stehen:

(...) it does follow from the fact that E_1 caused E_2 in virtue of a property f of E_1 and property g of E_2 that E_1 and E_2 are in lawlike connection partly or wholly in virtue of properties f and g . (...) it *does* follow from any statement that the event of the pears' being put on the scale caused the pointer to move to the two-pound mark, and the statement that it did so in virtue of only certain properties, that the events were in lawlike connection by way of those properties. We can call this the Principle of the Nomological Character of Causally-Relevant Properties. It is consistent with and indeed required by any tolerable account of causation and is integral to any account which takes causal relations precisely specified to *be* a species of lawlike relations.¹³³

Wenn also ein Ereignis E_1 kraft einer Eigenschaft f ein Ereignis E_2 aufgrund einer Eigenschaft g verursacht, dann stehen E_1 und E_2 – vollständig oder zumindest zum Teil – vermöge der Eigenschaften f und g in einer gesetzesartigen Verbindung zueinander. In Applikation auf das Birnen-Beispiel gilt dementsprechend: Stimmt man der Aussage zu, dass das Ereignis des Legens der Birnen in die Waagschale den Ausschlag des Zeigers der Waage bis zur Zweipfund-Marke verursacht hat, und akzeptiert zugleich auch die Behauptung, dass dies nur aufgrund gewisser Eigenschaften der beiden Ereignisse geschehen ist, dann folgt daraus, dass die beiden Ereignisse nur wegen dieser Eigenschaften in einer gesetzesartigen Beziehung zueinander stehen. Verallgemeinert man diese Annahme erhält man das Prinzip vom nomologischen Charakter kausal relevanter Eigenschaften (PNCKRE). Dieses Prinzip ist nicht nur verträglich mit allen zulässigen Vorstellungen über Kausalität, sondern wird von jenen sogar vorausgesetzt, so Honderich. Schließlich ist PNCKRE ein essentielles Element jeglicher Kausalitätstheorie, die Kausalrelationen als eine echte Teilmenge von gesetzesartigen Beziehungen auffasst.

Für den Fall, dass man PNCKRE zustimmt, stellt sich hinsichtlich mentaler Verursachung folgende Frage: „If a mental event causes a physical event, what is the causally-relevant prop-

¹³³ Honderich, Ted: „The Argument for Anomalous Monism“ in: *Analysis*, Vol. 42, 1982, S. 62.

erty, or what are the causally-relevant properties, of the mental event?“¹³⁴ Wenn ein mentales Ereignis also ein physikalisches Ereignis verursacht, was sind dann die kausal relevanten Eigenschaften des mentalen Ursacheereignisses? Da jedes mentale Ereignis nach Davidson mit einem physikalischen Ereignis identisch sein muss und da ein Ereignis qua Beschreibung mental oder physikalisch ist, muss jedes mentale Ereignis zwangsläufig auch eine physikalische Beschreibung und ergo auch physikalische Eigenschaften besitzen. Unter diesen Umständen stellt sich die Frage nach den tatsächlich kausal relevanten Eigenschaften natürlich umso dringlicher: sind die mentalen oder die physikalischen Eigenschaften des Ereignisses kausal relevant? Zur Beantwortung dieser Frage hält es Honderich für lohnenswert nochmals Davidsons Bemerkungen über Kausalität zu studieren, wobei Honderich sein Hauptaugenmerk auf zwei Punkte legt: Erstens stellt Kausalität für Davidson eine Beziehung zwischen einzelnen Ereignissen unabhängig von deren Art der Beschreibung dar. Zweitens befasst sich das Prinzip der kausalen Interaktion (PKI) laut Davidson mit den Extensionen von Ereignissen und lässt daher die mental-physikalische Unterscheidung außen vor. Nach Honderich bringt die erste Bemerkung die wahre Aussage zum Ausdruck, dass man über die Kausalbeziehung zwischen zwei Ereignissen sprechen kann, unabhängig davon wie man diese beschreibt. Tatsächlich kann man nach Honderich noch einen Schritt weiter gehen und Kausalrelationen präzisieren, indem man feststellt welche Eigenschaften der Ereignisse für die jeweilige Verursachungsbeziehung relevant sind. Über die Interpretation von Davidsons zweiter Bemerkung spekuliert Honderich wie folgt: Sollte sie so verstanden werden, dass das mentale Ereignis nicht aufgrund von gewissen Eigenschaften Ursache ist, dann ist die Aussage sicherlich falsch. Denn eine Unterscheidung zwischen kausal relevanten und kausal nicht relevanten Eigenschaften ist, wie etwa auch das Birnen-Beispiel zeigt, stets angeraten und es ist in keiner Weise ersichtlich, warum mentale Ereignisse eine Ausnahme von dieser Regel darstellen sollten. Eine zweite Deutungsmöglichkeit besteht darin, dass das mentale Ursacheereignis als mentales Ereignis das physikalische Wirkungsereignis verursacht hat. Damit würde gleichzeitig auch der Epiphanomalismus-Vorwurf gegenüber dem Mentalen ausgeräumt werden. Geht man allerdings davon aus, dass das physikalische Wirkungsereignis als physikalisches Ereignis eine Wirkung ist und akzeptiert zugleich PNCKRE, dann ist man letztlich gezwungen PAM zurückzuweisen. Verursacht nämlich das mentale Ursacheereignis als mentales Ereignis das physikalische Wirkungsereignis, dann sind die mentalen Eigenschaften des Ursacheereignisses letzten Endes kausal relevant für die Verursachung des physikalischen Ereignisses. Da aber kausal relevante Eigenschaften nach PNCKRE einen nomologischen Charakter

¹³⁴ Honderich, Ted: ‚*The Argument for Anomalous Monism*‘ in: *Analysis*, Vol. 42, 1982, S. 62.

ter besitzen, müsste es folglich auch strikte psychophysische Gesetze geben. Will man hingegen an PAM festhalten und insistiert indes auch auf der Wahrheit von PNCKRE, dann wird man dadurch dazu genötigt die Vorstellung aufzugeben, dass das mentale Ereignis als *mentales* Ereignis die physikalische Wirkung verursacht hat. Demnach kommt als dritte Möglichkeit die Frage nach den kausal relevanten Eigenschaften zu beantworten die These in Frage, dass das mentale Ereignis als *physikalisches* Ereignis die Handlung verursacht hat. Der Entschluss die Frage auf diese Weise zu beantworten ist jedoch folgeschwer – insofern, als dass PKI dann in einem völlig neuen Licht erscheint. Aus der Behauptung, dass das Mentale kausal mit dem Physikalischen interagiert, wird die Aussage, dass das Mentale als Physikalisches mit dem Physikalischen in kausale Wechselwirkung tritt. Dies widerspricht jedoch offensichtlich der Überzeugung, die ursprünglich hinter PKI stand – und zwar dem Glauben an die kausale Wirksamkeit des Mentalen. Diesem Grundsatz zufolge ist ein Ereignis als etwas Mentales ein nicht reduzierbarer Teil einer vollständigen Erklärung von Handlungen. Der mentale Aspekt eines Ereignisses darf diesem Gedanken zufolge also nicht aus einer umfassenden Erklärung einer Handlung eliminierbar sein. Genau dieser Punkt rechtfertigt auch das Zurückweisen des Epiphänomenalismus-Vorwurfs; schließlich wird auf diese Weise eine gewisse Form von Eigenständigkeit des Mentalen gewahrt.¹³⁵

Anschließend analysiert Honderich, ob es Wege gibt diese Vorstellung vom Mentalen fruchtbar zu vertreten. Etwa indem man behauptet das Ereignis wäre nicht die Ursache gewesen, die es war, wenn es nicht gleichzeitig sowohl ein mentales als auch ein physikalisches Ereignis gewesen wäre. Dadurch würde dem Mentalen jedoch sicherlich keine kausale Relevanz in Bezug auf die Handlung verliehen, so Honderich. Insofern kann auch mit dieser Argumentation nicht die kausale Wirksamkeit des Mentalen gewährleistet werden. Diese wird schließlich auch nicht von der Tatsache garantiert, dass man von einem Ereignis sagen kann es ist eine Ursache, selbst wenn es anhand von kausal irrelevanten Eigenschaften beschrieben wird. Eine weitere verbleibende Möglichkeit beschreibt Honderich so:

The doctrine of supervenience, so far unmentioned, will come to mind. The picture we get then is this: it is a mental event as physical which causes an action; lawlike connection holds between the mental event as physical and the action, but not between the mental event as mental and that same event as physical; however, since the event as mental supervenes on the event as physical, the event as mental is efficacious with respect to the action. The final claim turns on what supervenience comes to, and what it comes to cannot be lawlike connection between the mental and the physical, and is indeed to be understood as no more than the holding of certain universal material conditionals. That is to say that the connection between the mental and the physical is accidental. There is no nomic necessity about the event as physical being the mental event it is. Here, it seems, we do not get the efficacy of the mental.¹³⁶

¹³⁵ Vgl. Honderich, Ted: ‚*The Argument for Anomalous Monism*‘ in: *Analysis*, Vol. 42, 1982, S. 62ff.

¹³⁶ Honderich, Ted: ‚*The Argument for Anomalous Monism*‘ in: *Analysis*, Vol. 42, 1982, S. 64.

Zunächst hat es den Anschein als könnte eine Supervenienztheorie dem oben skizzierten Bild des Mentalen eine gewisse Stimmigkeit verleihen, so Honderich. Einem solchen Ansatz gemäß stellt sich die Situation in etwa folgendermaßen dar: das mentale Ereignis verursacht als physikalisches Ereignis eine Handlung. Die gesetzesartige Verbindung besteht dabei zwischen dem mentalen Ereignis als physikalischem Ereignis und der Handlung und nicht etwa zwischen dem mentalen Ereignis als mentalem Ereignis und demselben Ereignis als physikalischem Ereignis. Da aber das Ereignis als mentales Ereignis auf dem Ereignis als physikalischem Ereignis superveniert, wird dadurch auch die kausale Wirksamkeit des Ereignisses als mentales Ereignis hinsichtlich der Handlung abgesichert. Allerdings kann auch eine Supervenienztheorie keine gesetzesartige Verbindung zwischen Mentalen und Physikalischen etablieren. Somit ist die Beziehung zwischen Mentalem und Physikalischem rein zufällig. Für ein beliebiges physikalisches Ereignis gibt es folglich keine nomologische Notwendigkeit das mentale Ereignis zu sein, das es de facto ist. Demnach bürden auch Supervenienztheorien nicht für eine kausale Wirksamkeit des Mentalen.

Schlussendlich resümiert Honderich die Zielsetzung seines Aufsatzes: „I have not looked into the question of the truth of the first and third claims out of which Anomalous Monism arises. All I tried to establish is that the three claims, together with the fact of causally-relevant properties, the principle about their nomological character, and the conviction of the efficacy of the mental, are bad news for Anomalous Monism.“¹³⁷ Laut Honderich entsteht für Vertreter des Anomalen Monismus ein Problem, wenn sie zusätzlich zu den drei Prämissen PKI, PNCK und PAM, dem Prinzip vom nomologischen Charakter kausal relevanter Eigenschaften (PNCKRE), zustimmen und auf der Wirksamkeit des Mentalen beharren. Falls nämlich ein mentales Ereignis eine Handlung (i.e. ein physikalisches Ereignis) verursacht, stellt sich für Honderich die Frage welche Eigenschaften des Ursacheereignisses kausal relevant waren – die mentalen oder die physikalischen.¹³⁸ Nun scheint es jedoch keine Möglichkeit zu geben, dass das mentale Ereignis als mentales Ereignis die Ursache der Handlung war – ansonsten müsste es nämlich nach PCKNRE in einer gesetzesartigen Beziehung zu einem physikalischen Ereignis stehen. Genau dies wird aber von PAM ausgeschlossen. Insofern muss das mentale Ereignis als physikalisches Ereignis die Handlung verursacht haben. Dann sind allerdings die mentalen Eigenschaften des Ursacheereignisses nicht kausal relevant für das Eintreten des Wirkungereignisses. Wenn dies der Fall ist, scheint dadurch die kausale Wirksamkeit

¹³⁷ Honderich, Ted: ‚*The Argument for Anomalous Monism*‘ in: *Analysis*, Vol. 42, 1982, S. 62.

¹³⁸ Da Davidson zufolge jedes mentale Ereignis mit einem physikalischen Ereignis identisch sein muss, besitzt jedes mentale Ereignis neben seiner mentalen Beschreibung automatisch auch eine physikalische Beschreibung. Das Haben einer physikalischen bzw. mentalen Beschreibung ist für ein Ereignis laut Honderich äquivalent damit physikalische bzw. mentale Eigenschaften zu besitzen.

des Mentalen akut bedroht zu sein. Ein Verfechter des Anomalen Monismus ist nämlich weiterhin verpflichtet von der Wahrheit von PKI, PNCK und PAM auszugehen – schließlich leitet Davidson die Gültigkeit des Anomalen Monismus aus diesen drei Prämissen ab. Zudem zeige das Birnen-Beispiel, dass in Kausalbeziehungen zwischen einzelnen Ereignissen einige Eigenschaften kausale Relevanz besitzen, andere nicht. Da dieses Prinzip offensichtlich leicht zu verallgemeinern ist, folgt daraus die Wahrheit von PNCKRE. Ist man schließlich geneigt von der kausalen Wirksamkeit des Mentalen auszugehen – was sicherlich eine Grundmotivation hinter PKI ist – dann gerät man spätestens an dieser Stelle in die missliche Lage entweder PAM, PNCKRE oder die kausale Wirksamkeit des Mentalen (KWM) aufzugeben. Denn diese drei Punkte können nicht miteinander in Einklang gebracht werden. Für den Fall, dass es strikte psychophysische Gesetze gibt, sind zwar PNCKRE und KWM wahr, PAM wird dadurch aber offensichtlich falsifiziert. Wenn wiederum PAM und KWM der Wahrheit entsprechen, dann müsste PNCKRE falsch sein – schließlich besäßen in diesem Fall die vermeintlich kausal wirksamen mentalen Eigenschaften keine kausale Relevanz; da ein beliebiges mentales Ursacheereignis nicht als mentales, sondern ausschließlich als physikalisches Ereignis in einer gesetzesartigen Verbindung zum physikalischen Wirkungsereignis stünde. Insofern würden dann auch nur die physikalischen Eigenschaften der beiden Ereignisse in einer gesetzesartigen Verbindung zueinander stehen und hätten somit den Anspruch auf kausale Relevanz. Geht man schließlich von der Wahrheit von PAM und PNCKRE aus, so folgt daraus die kausale Wirkungslosigkeit des Mentalen, also die Falschheit von KWM. Denn falls es keine strikten psychophysischen Gesetze gibt und kausal relevante Eigenschaften aber dennoch in einer gesetzesartigen Verbindung zueinander stehen, kommen in Fällen mentaler Verursachung nur die physikalischen Eigenschaften als kausal relevante Eigenschaften in Frage. Besitzt aber ein mentales Ereignis nur als physikalisches Ereignis kausale Relevanz, dann scheint es dadurch jeglicher kausaler Wirksamkeit beraubt zu sein.

Kontrafaktische Abhängigkeit als Lösung für das Anomalie-Problem

In ihrem Aufsatz *Mind Matters* verteidigen Ernest Le Pore und Barry Loewer den Anomalen Monismus gegen Ted Honderichs Epiphenomenalismus-Vorwurf. Dazu bilden sie zunächst den allgemeinen Hintergrund des Problems mentaler Verursachung ab und stellen anschließend Davidsons Lösungsversuch dieses Problems jeweils anhand einiger Thesen dar:

- 1.) The mental and the physical are distinct.
- 2.) The mental and the physical causally interact.
- 3.) The physical is causally closed.

(...)

3'.) All causation is reducible to, or grounded in, physical causation.

(...)

Our primary concern here is how Davidson's account of the relation between the mental and the physical, which he calls "anomalous monism" (AM), attempts to resolve the dilemma. AM consists of the following three theses:

- 4.) There are no strict psychophysical or psychological laws and in fact all strict laws are expressed in a purely physical vocabulary (the anomalousness of the mental).
- 5.) Mental events causally interact with physical events.
- 6.) Event *c* causes event *e* only if there is a strict causal law which subsumes *c* and *e* (entails that *c* causes *e*) (the nomological character of causality).¹³⁹

Die Prinzipien 1.) bis 3'.) drücken ein Dilemma hinsichtlich mentaler Verursachung aus, denn jede These für sich erscheint als durchaus einleuchtend, allerdings hat es den Anschein als könne man alle drei zusammen nicht kohärent vertreten. Dies wird zum Beispiel offensichtlich, wenn man mit diesen drei Aussagen auf Ereignisse Bezug nimmt – dann würde nämlich gelten: erstens ist kein mentales Ereignis ein physikalisches Ereignis, zweitens verursachen einige mentale Ereignisse physikalische Ereignisse (und umgekehrt) und drittens müssen alle Ursachen von physikalischen Ereignissen selbst physikalische Ereignisse sein. Der Widerspruch ist offensichtlich. Davidsons Anomaler Monismus ist ein Versuch die Beziehung zwischen Mentalem und Physikalischem adäquat auszudrücken und gleichzeitig die oben skizzierte Unvereinbarkeit aufzuheben. Zu diesem Zweck schlägt Davidsons folgenden Weg ein: 4.) kann als eine Variante von 1.) betrachtet werden, denn aus 4.) folgt, dass mentale Eigenschaften von physikalischen Eigenschaften verschieden sind. Gemeinhin wird nämlich eine Eigenschaft, die durch M ausgedrückt werden kann, als reduzierbar auf eine Eigenschaft, die durch P bezeichnet werden kann, erachtet, wenn es ausnahmslos gültige Brückengesetze gibt, die M und P in Beziehung zueinander setzen. Somit sichert also das Fehlen von strikten, psychophysischen Gesetzen die Eigenständigkeit mentaler Eigenschaften ab. Das Prinzip 6.) besagt wiederum, dass ein Ereignis *c* ein Ereignis *e* nur dann verursacht, wenn es eine singuläre Beschreibung *D* von *c* und eine singuläre Beschreibung *D'* von *e* sowie ein Gesetz *L* gibt, so dass gilt: sind *L* und „*D* hat sich ereignet“ der Fall, dann muss auch gelten „*D* verursacht *D'*“. Zusammen mit der Annahme, dass strikte Gesetze nur in physikalischem Vokabular formuliert werden können (vgl. 4.), folgt daraus, dass physikalische Ereignisse nur physikalische Ereignisse als Ursachen haben können oder in anderen Worten, dass die Verursachung aller Ereignisse im Grunde physikalisch sein muss. Insofern beinhaltet 6.) also die Aussage von 3.) bzw. 3'.). In diesem Zusammenhang postuliert Davidson zudem, dass strikte Gesetze keine *ceteris paribus* Klausel enthalten dürfen, sondern vielmehr eine ausnahmslose Gültigkeit besitzen müssen. Dementsprechend können psychologische und psychophysische Gesetze ihrem

¹³⁹Le Pore, Ernest & Loewer, Barry: „*Mind Matters*“ in: *The Journal of Philosophy*, Vol. 84, 1987, S. 630f.

Wesen nach nicht strikt (d.h. ausnahmslos gültig) sein, sondern gelten jeweils nur unter gleichbleibenden Umständen. Da die Prinzipien 1.) bis 3'.) einerseits und 4.) bis 6.) andererseits deutliche Parallelen aufweisen, stellt sich die Frage, ob letztere auch miteinander unvereinbar sind. Darauf geben Le Pore und Loewer eine deutliche Antwort:

(...) [T]hey are not. Davidson shows that they all can be true if (and only if) mental events are identical to physical events (...). Let us say that an event *e* is a physical event just in case *e* satisfies a basic physical predicate (that is, a physical predicate appearing in a strict law). Since only physical predicates (or predicates expressing properties reducible to basic physical properties) appear in strict laws, it follows that every event that enters into causal relations satisfies a basic physical predicate. So, those mental events which enter into causal relations are also physical events.

AM is committed only to a partial endorsement of (1). The mental and physical are distinct insofar as they are not linked by strict law – mental properties are not reducible to physical properties – but they are not distinct insofar as mental events are physical events. This being so, one might wonder whether AM also only partially endorses claims (2) and (3). In fact, Davidson's views have been criticized precisely on the point of (2).¹⁴⁰

Obwohl also die Sätze 4.) – 6.) inhaltlich frappierende Ähnlichkeiten mit den Behauptungen 1.) – 3'.) aufweisen, können erstere im Gegensatz zu zweiten miteinander in Einklang gebracht werden, so Le Pore und Loewer. Diese Harmonie erreicht Davidson, indem er ein Identitätsverhältnis zwischen mentalen und physikalischen Ereignissen postuliert. Dabei ist ein Ereignis *e* der Klasse physikalischer Ereignisse zuzurechnen, wenn es einem basalen physikalischen Prädikat – i.e. ein physikalisches Prädikat, das in einem strikten Gesetz vorkommt – genügt. Da ausschließlich basale physikalische Prädikate in Kausalbeziehungen eintreten können, müssen mentale Ereignisse, die ein Element einer Kausalrelation sind, ergo gleichzeitig auch physikalische Ereignisse sein. Die besondere Komponente des Anomalen Monismus, die eine Auflösung des durch 1.) – 3'.) beschriebenen Dilemmas ermöglicht, liegt in einem speziellen Verständnis von 1.) So sieht sich der Anomale Monismus der Aussage von 1.) gegenüber nur teilweise verpflichtet. Demnach sind das Mentale und das Physikalische also insofern unterschiedlich, als dass sie nicht mittels eines strikten Gesetzes miteinander verknüpft werden können – das bedeutet, dass mentale Eigenschaften nicht reduzierbar auf und daher auch nicht identisch mit physikalischen Eigenschaften sind. Allerdings sind das Mentale und das Physikalische insofern nicht unterschiedlich, als dass mentale Ereignisse (qua Beschreibung) zugleich auch physikalische Ereignisse sind. Da sich der Anomale Monismus nicht zur Gänze zum Prinzip 1.) bekennt, wurde bisweilen gemutmaßt, dass er auch 2.) und 3.) nur partiell anerkennt. Insbesondere im Hinblick auf 2.) ist Davidsons Theorie vehement angefochten worden. Genau auf diesen Punkt zielt auch Honderichs Kritik am Anomalen Monismus ab. Die Kernaussage seiner Missbilligung gegenüber der Davidsonschen Konzeption besteht – wie oben bereits ausführlich dargelegt – in der Behauptung, dass das Mentale

¹⁴⁰ Le Pore, Ernest & Loewer, Barry: „*Mind Matters*“ in: *The Journal of Philosophy*, Vol. 84, 1987, S. 632.

nicht qua Mentalem verursacht, sondern nur qua Physikalischem. Anders ausgedrückt prangert Honderich den Anomalen Monismus als eine im Grunde epiphänomenalistische Position an, denn die gesamte kausale Arbeit werde schließlich vom Ursacheereignis als physikalischem Ereignis geleistet. Auf die Entgegnung Davidsons zu diesem Vorwurf, dass Verursachung eine Relation zwischen Ereignissen als solchen ist und nicht eine Beziehung zwischen Ereignissen als mental oder physikalisch verstanden werden kann, wenden die Kritiker seiner Theorie meist ein, die Annahme, dass einige Eigenschaften eines Ereignisses (im Gegensatz zu anderen) kausal relevant sind, sei durchaus sinnvoll. Sobald man der Einteilung in kausal relevante und kausal nicht relevante Eigenschaften eines Ereignisses zustimmt – und eine solche Unterscheidung ist scheinbar allein schon durch den gesunden Menschenverstand geboten – scheint die kausale Irrelevanz mentaler Charakteristika besiegelt zu sein. Le Pore und Loewer rekonstruieren das zugrundeliegende Argument folgendermaßen:

The argument is that, since, according to AM, *c* causes *e* only if there is a strict law that subsumes *c* and *e* and since strict laws contain only physical (never mental) predicates, it follows that the mental features of events *c* and *e* are irrelevant to whether they are causally connected. The physical features of the event suffice to fix, given the strict laws, *all* causal connections. Mental features neither suffice nor are required to fix causal connections. The argument is powerful. The conclusion the authors draw from it is that on AM *the mind does not matter*; that a neural event has a certain intentional content is as irrelevant to its effect as the fact that the sounds are meaningful is to the sounds causing the glass to break.¹⁴¹

Das Argument beruft sich also maßgeblich darauf, dass gemäß des Anomalen Monismus *c* nur dann *e* verursachen kann, wenn es ein striktes Gesetz gibt, unter das *c* und *e* fallen. Da strikte Gesetze aber ausschließlich auf eine rein physikalische Lexik zurückgreifen, folge daraus, dass die mentalen Merkmale der Ereignisse *c* und *e* keine Rolle für deren kausale Verbindung spielen. Denn schon allein die physikalischen Charakteristika von *c* und *e* seien hinreichend, um die Kausalbeziehung zwischen den beiden Ereignissen gänzlich festzulegen; die mentalen Attribute hingegen seien dafür weder hinreichend noch notwendig. Die Schlussfolgerung, die Verfechter dieses Gedankengangs daraus ziehen, besagt, dass das Mentale im Anomalen Monismus nicht von Belang sei. Le Pores und Loewers Analyse dieses Einwands gegen den Anomalen Monismus gelangt jedoch zu einem gegenteiligen Ergebnis:

But is this criticism of AM correct? We claim that it is not, and that it rests on a simple, but perhaps not obvious, confusion. The confusion is between two ways in which properties of an event *c* may be said to be causally relevant and irrelevant. Consider the following locutions:

- a) Properties *F* and *G* are relevant₁ to making it the case that *c* causes *e*,
and
- b) *c*'s possessing property *F* is causally relevant₂ to *e*'s possessing property *G*.

¹⁴¹ Le Pore, Ernest & Loewer, Barry: „*Mind Matters*“ in: *The Journal of Philosophy*, Vol. 84, 1987, S. 634.

We will say that (a) holds iff c has F and e has G , and there is a strict law that entails F s cause G s. It is in this sense that it is c 's having F and e 's having G "make it the case" that c causes e . Relevance₂ is a relation among c , one of its properties F , e , and one of its properties G . It holds when c 's being F brings about that e is G . We shall argue that those who charge AM with epiphenomenalism are guilty of confusing relevance₁ with relevance₂.¹⁴²

Der Epiphänomenalismus-Vorwurf gegen den Anomalen Monismus ist laut Le Pore und Loewer nicht gerechtfertigt, weil die Position zwei unterschiedliche Auffassungen von kausaler Relevanz miteinander vermischt. Demnach gibt es zwei Möglichkeiten wie die Eigenschaft F eines Ereignisses c kausal relevant für die Eigenschaft G eines Ereignisses e sein kann. Diese beiden Arten kausaler Relevanz lassen sich anhand der folgenden beiden Sätze voneinander unterscheiden:

- a.) Die Eigenschaften F und G sind relevant₁ dafür, dass es der Fall ist, dass c e verursacht.
- b.) c 's Haben der Eigenschaft F ist kausal relevant₂ für e 's Besitzen der Eigenschaft G .

Die Aussage a.) gilt genau dann, wenn c F und e G hat und es ein striktes Gesetz gibt, das es nach sich zieht, dass F s G s verursachen. In diesem Sinne sind c 's Haben von F und e 's Besitzen von G dafür verantwortlich, dass c e verursacht. Im Gegensatz dazu ist Satz b.) genau dann gültig, wenn die Tatsache, dass c F hat bewirkt, dass e G hat. Denjenigen, die den Anomalen Monismus des Epiphänomenalismus' verdächtigen, entgegnen Le Pore und Loewer Relevanz₁ und Relevanz₂ nicht zu unterscheiden: „The heart of our response to the claim that AM is committed to epiphenomenalism is this: AM entails that mental features are causally irrelevant₁, but does not entail that they are causally irrelevant₂.“¹⁴³ Während Le Pore und Loewer also der Aussage zustimmen würden, dass mentale Charakteristika dem Anomalen Monismus zufolge kausal irrelevant₁ sind, würden sie hingegen jedoch negieren, dass sie auch kausal irrelevant₂ sind. Schließlich repräsentiert kausale Irrelevanz₂ lediglich eine Art kontrafaktischer Abhängigkeit zwischen den Eigenschaften F von c und G von e . Als Test für kausale Relevanz₂ schlagen Le Pore und Loewer folgende Vorgehensweise vor:

- (I) c 's being F is causally relevant₂ to e 's being G iff
 - i. c causes e .
 - ii. Fc and Ge .
 - iii. $\neg Fc > \neg Ge$.
 - iv. Fc and Ge are logically and metaphysically independent.¹⁴⁴

¹⁴² Le Pore, Ernest & Loewer, Barry: „Mind Matters“ in: *The Journal of Philosophy*, Vol. 84, 1987, S. 634f.

¹⁴³ Le Pore, Ernest & Loewer, Barry: „Mind Matters“ in: *The Journal of Philosophy*, Vol. 84, 1987, S. 635f.

¹⁴⁴ Le Pore, Ernest & Loewer, Barry: „Mind Matters“ in: *The Journal of Philosophy*, Vol. 84, 1987, S. 635.

Die Tatsache, dass c F hat ist demgemäß also genau dann kausal relevant₂ für e's Haben von G, wenn folgende vier Bedingungen erfüllt sind: Erstens ist c die Ursache von e. Zweitens besitzt das Ereignis c die Eigenschaft F und das Ereignis e instanziiert die Eigenschaft G. Drittens gilt, wenn c nicht F hat, dann hat e auch nicht G. Viertens ist c's Haben von F sowohl logisch als auch metaphysisch unabhängig von e's Besitzen von G. Diese Überlegungen legen den Schluss nahe, dass mentale Merkmale in diesem Sinne durchaus kausale Relevanz₂ besitzen können. Im Gegensatz dazu ist es ebenso sehr evident, dass das Mentale gemäß des Anomalen Monismus' niemals kausale Relevanz₁ besitzen kann. Schließlich müsste es ansonsten strikte Gesetze geben, die mentale Eigenschaften mit physikalischen Eigenschaften in eine kausale Beziehung setzen würden. Da aber einerseits der Anomale Monismus die Möglichkeit strikter psychophysischer Gesetze a priori ausschließt, andererseits jedoch die Definition kausaler Relevanz₁ die Verkettung von kausal relevanten₁ Eigenschaften über strikte Gesetze einfordert, ist somit die kausale Irrelevanz₁ mentaler Eigenschaften im Anomalen Monismus festgeschrieben. Verzichtet man bei der Definition kausaler Relevanz allerdings auf die Bedingung, dass die entsprechenden Eigenschaften mittels eines strikten Gesetzes aneinander gekoppelt sind – so wie dies bei kausaler Relevanz₂ der Fall ist – dann können mentale Eigenschaften auch im Rahmen des Anomalen Monismus kausal relevant₂ sein. Akzeptiert man diesen Gedankengang, so bleiben laut Le Pore und Loewer dennoch zwei Fragen offen: erstens, ob die kausale Irrelevanz₁ mentaler Eigenschaften bereits Honderichs Epiphenomenalismus-Vorwurf gegenüber dem Anomalen Monismus rechtfertigt. Und zweitens, ob es weitere Bedingungen für den Test kausaler Relevanz₂ gibt, die sobald sie zu (I) hinzugefügt sind, zur Folge haben, dass im Anomalen Monismus das Mentale auch kausal irrelevant₂ wird.¹⁴⁵

Im Hinblick auf die erste Frage versucht Honderich seine Antwort durch das Prinzip des nomologischen Charakters kausal relevanter Eigenschaften (PNCKRE) zu untermauern. Dabei geben Le Pore und Loewer jedoch zu bedenken: „If Honderich intends for the principle of nomological relevance to include nonstrict as well as strict laws, then AM is compatible with causal relevance (in Honderich's sense) of psychological properties. If he intends for the principle to include only strict laws, then it is an unacceptable principle. It is implausible that there are any strict laws linking 'is a match striking' with 'is a match lighting'. So, on the strict law construal of Honderich's principle, being a match striking is not causally relevant to the match's lighting. On this construal, Honderich's principle would render virtually all properties of events causally irrelevant₂. This certainly seems wrong.“¹⁴⁶ Laut Le Pore und Loewer

¹⁴⁵ Vgl. Le Pore, Ernest & Loewer, Barry: „Mind Matters“ in: *The Journal of Philosophy*, Vol. 84, 1987, S. 636f.

¹⁴⁶ Le Pore, Ernest & Loewer, Barry: „Mind Matters“ in: *The Journal of Philosophy*, Vol. 84, 1987, S. 637f.

ist eine Frage von entscheidender Bedeutung für PNCKRE – und zwar, ob kausal relevante Eigenschaften ausschließlich mittels eines strikten Gesetzes miteinander verbunden sein müssen oder nicht. Wenn dies nämlich nicht zwangsläufig der Fall sein muss, dann könnten auch psychologische Eigenschaften kausal relevant für physikalische Eigenschaften sein. Schließlich ist laut Davidson die Existenz psychophysischer und psychologischer Kausalgesetze durchaus möglich – seine einzige Einschränkung in diesem Zusammenhang besteht darin, dass diese psychophysischen Gesetze niemals strikt sein können. Wäre also gemäß PNCKRE eine Verbindung kausal relevanter Eigenschaften über ein nicht-striktes Gesetz möglich, dann könnten auch mentale Eigenschaften kausale Relevanz in Honderichs Sinne besitzen – dementsprechend würde auch seinem Epiphänomenalismus-Vorwurf gegenüber dem Anomalen Monismus jegliche Grundlage entzogen. Falls Honderich aber ausschließlich strikte Gesetze zuließe, um den Zusammenhang von kausal relevanten Eigenschaften getreu PNCKRE zu erklären, so könnten mentale Eigenschaften tatsächlich keine kausale Relevanz in diesem Sinne besitzen und wären letzten Endes rein epiphänomenal. Allerdings wäre es dann auch ein Leichtes PNCKRE aufgrund eben dieser strengen Auslegung abzulehnen. Schließlich scheint die Existenz eines strikten Kausalgesetzes, das beispielsweise die Eigenschaft ‚ist ein Streichholz-Reiben‘ mit der Eigenschaft ‚ist ein Streichholz-Entzünden‘ in Beziehung setzt, als äußerst unplausibel. Wenn jedoch nur Eigenschaften, die mittels eines strikten Gesetzes verknüpft sind, kausal relevant füreinander sein können, dann könnte demzufolge auch die Eigenschaft ‚ist ein Streichholz-Reiben‘ niemals kausal relevant für die Eigenschaft ‚ist ein Streichholz-Entzünden‘ sein. Dies ist nicht nur im höchsten Maße kontraintuitiv, sondern hätte auch zur Folge, dass praktisch alle Eigenschaften eines Ereignisses kausal irrelevant₂ wären. Alles in allem kann also ein Fehlen kausaler Relevanz₁ von mentalen Eigenschaften im Anomalen Monismus nicht Honderichs Epiphänomenalismus-Vorwurf gegenüber dieser Position rechtfertigen. Die entscheidende Frage ist nämlich, ob Honderich in kausaler Relevanz₁ die einzig adäquate Form kausaler Relevanz überhaupt sieht. Sollte dies der Fall sein, so wäre seine Auffassung kausaler Relevanz jedoch eindeutig zu eng gefasst; schließlich würden in diesem Fall nahezu alle Eigenschaften eines Ereignisses ihre kausale Relevanz verlieren. Sollte Honderich jedoch eine Interpretation von kausaler Relevanz im Sinne von kausaler Relevanz₂ zulassen, dann würden mentale Eigenschaften im Anomalen Monismus nicht von seinem Epiphänomenalismus-Vorwurf tangiert werden. Denn der Verzicht auf die Prämisse, dass kausal relevante Eigenschaften über ein striktes Gesetz miteinander verknüpft sein müssen, reduziert die Bedingung kausaler Relevanz lediglich auf eine kontrafaktische Abhängig-

keit. Eine solche wäre allerdings im Falle mentaler Eigenschaften im Anomalen Monismus durchaus gegeben.

Eine andere Möglichkeit die Kritiker des Anomalen Monismus verwenden könnten, um die kausale Relevanz mentaler Eigenschaften im Rahmen dieses Ansatzes anzuzweifeln, besteht laut Le Pore und Loewer in einer Verschärfung der hinreichenden Bedingungen für kausale Relevanz₂. Denkbar wären etwa die folgenden beiden Zusatzbedingungen:

- (II) c 's being F is causally irrelevant₂ to e 's being G , if there is a property F^* of c such that $(F^*c \ \& \ -Fc) > Ge$ holds nonvacuously.
(...)
- (III) c 's being F is relevant₂ to e 's being G iff the conditions in (I) are satisfied and there is *no* property F^* of c such that $(F^*c \ \& \ -Fc) > Ge$ holds nonvacuously.¹⁴⁷

Genauer genommen handelt es sich nicht um zwei voneinander unabhängige Zusatzbedingungen für kausale Relevanz₂, sondern (II) ist eine hinreichende Bedingung für kausale Irrelevanz₂, die mit (III) in eine notwendige Bedingung kausaler Relevanz₂ überführt wird. Dabei besagt (II), dass c 's Besitzen von F kausal irrelevant₂ für e 's Haben von G ist, wenn es eine Eigenschaft F^* von c gibt, von der gilt: das Konditional $F^*c \ \& \ -Fc \rightarrow Ge$ ist zumindest für ein c wahr. Das bedeutet, es existiert mindestens ein c , das F^* hat, zugleich aber nicht F besitzt, nichtsdestotrotz aber die Ursache dafür ist, dass e G instanziiert. Daraus leitet (III) eine positive Bedingung für kausale Relevanz₂ ab: c 's Haben von F ist genau dann kausal relevant₂ für e 's Besitzen von G , wenn die Bedingungen von (I) erfüllt sind und c gleichzeitig keine Eigenschaft F^* hat, so dass $F^*c \ \& \ -Fc \rightarrow Ge$ gilt. Gäbe es nämlich eine solche Eigenschaft F^* , die c besitzt und die selbst wenn c nicht F hat dafür sorgt, dass e G besitzt, dann würde F^* F sozusagen von dessen kausaler Relevanz₂ für e 's Haben von G abtrennen. Übertragen auf das Problem mentaler Verursachung im Rahmen des Anomalen Monismus könnte ein Verfechter von (III) behaupten, dass z.B. die neuronalen Eigenschaften (Nc) eines Ursacheereignisses dessen mentale Eigenschaften (Mc) von den behavioralen Eigenschaften des Wirkungsereignisses (Be) abtrennen. In anderen Worten besitzen die mentalen Eigenschaften eines Ursacheereignisses demgemäß keinerlei kausale Relevanz₂ für eventuelle Wirkungen im Verhalten. Ein Beispiel: Wenn ich Durst verspüre und meinen Arm nach einem Glas Wasser ausstrecke, dann würden nach diesem Bild die neuronalen Eigenschaften des Ursacheereignisses dessen intentionale (i.e. mentale) Eigenschaften abhalten kausal relevant₂ für das Strecken meines Armes zu sein. Diese Schlussfolgerung lässt sich einigen Kritikern zufolge unmittelbar aus dem Anomalen Monismus ziehen. Letztendlich wird diese Konklusion damit begründet, dass es strikte Gesetze gebe, die neuronale Eigenschaften mit behavioralen Eigen-

¹⁴⁷ Le Pore, Ernest & Loewer, Barry: „*Mind Matters*“ in: *The Journal of Philosophy*, Vol. 84, 1987, S. 638.

schaften in Beziehung setzen. Le Pore und Loewer bezweifeln allerdings, dass es strikte Gesetze gibt, die neuronale Eigenschaften mit mentalen Eigenschaften verknüpfen. Somit stehen sie auch der These, es *folge* aus dem Anomalen Monismus, dass neuronale Eigenschaften mentale Eigenschaften von deren kausaler Relevanz abtrennen, skeptisch gegenüber. Es scheint jedoch ein Faktum zu sein, dass in dem oben beschriebenen Szenario die neuronalen Eigenschaften in der Tat die mentalen Eigenschaften von jeglicher kausaler Relevanz abschirmen. Die Hauptsorge von Le Pore und Loewer besteht nun darin, dass die kausale Irrelevanz₂ mentaler Eigenschaften wirklich aus dem Anomalen Monismus folgt, wenn man an (II) als Bedingung kausaler Irrelevanz₂ festhält. Dazu merken sie an:

In response to this, notice first that (II)'s rendering the mental causally irrelevant₂ is independent of AM, at least to the extent that the problem-creating counterfactual, $(Nc \ \& \ -Mc) > Be$, holds whether or not there is a strict law linking N with B . So anyone who adopts (II) as a condition on causal irrelevance₂ will be committed to the causal irrelevance₂ of the mental in this case. But it seems to us that (II) is not a correct condition on irrelevance₂. It renders even properties connected by strict law causally irrelevant₂. To see this, consider the neural event c and the behavioral event e (...). c possesses basic physical property P and mental property M (being a desire to quench (...) thirst), and e possesses the property B (being a certain movement of the hand). Assuming a strict law between P and B , it follows that:

(S) $(-Mc \ \& \ Pc) > Be$.

So, P screens off M from B . Now consider the counterfactual:

(T) $(-Pc \ \& \ Mc) > Be$.

It can be shown that (T) is compatible with AM *and* (S). Furthermore, it is plausible that (T) is in fact true. If c had been a desire to quench thirst but had not been P , it would have had some other property P^* . Furthermore, c still would have resulted in an e that has the property B .¹⁴⁸

Nach Le Pore und Loewer wird das Mentale bereits durch die Zustimmung zu (II) – also unabhängig vom Anomalen Monismus – kausal irrelevant₂ gemacht. Zumindest in dem Ausmaß, als dass das Problem erzeugende kontrafaktische Konditional $(Nc \ \& \ -Mc) > Be$ selbst dann gültig ist, wenn N und B nicht durch ein striktes Gesetz verknüpft sind. Aus diesem Grund verpflichtet die Akzeptanz von (II) als Bedingung kausaler Irrelevanz₂ zu der Annahme, dass das Mentale generell kausal irrelevant₂ ist. Le Pore und Loewer argumentieren jedoch, dass (II) nicht den Status einer angemessenen Bedingung für Irrelevanz₂ innehat, weil (II) selbst Eigenschaften, die durch ein striktes Gesetz miteinander verbunden sind, kausal irrelevant₂ werden lässt. Dazu zitieren sie nochmals obiges Beispiel, in dem ein neuronales Ereignis c ein behaviorales Ereignis e verursacht; wobei c eine basale physikalische Eigenschaft P und eine mentale Eigenschaft M (ein Verlangen Durst zu stillen zu sein) besitzt, wohingegen e ein gewisse Eigenschaft B (ein bestimmte Handbewegung zu sein) hat. Für den Fall, dass ein striktes Gesetz existiert, das P und B miteinander verknüpft, gilt die kontrafaktische Bedingung (S) :

(S) $(-Mc \ \& \ Pc) > Be$

¹⁴⁸ Le Pore, Ernest & Loewer, Barry: „*Mind Matters*“ in: *The Journal of Philosophy*, Vol. 84, 1987, S. 639.

Denn ein Haben der Eigenschaft P eines Ursacheereignisses wäre in diesem Fall bereits offensichtlich allein hinreichend für das Eintreten eines Wirkungsereignisses, das die Eigenschaft B besitzt. D.h. auch wenn das Ursacheereignis nicht die Eigenschaft M gehabt hätte, so würde dessen ungeachtet das Wirkungsereignis die Eigenschaft B instanziiieren. In anderen Worten, wäre das Strecken meines Armes auch dann eingetreten, wenn ich kein Verlangen verspürt hätte meinen Durst zu stillen, sondern mein Gehirn lediglich die physikalische Eigenschaft P instanziiert hätte. Insofern schirmt P auch M von B ab. An diesem Punkt verweisen Le Pore und Loewer auf ein ähnliches kontrafaktisches Konditional, das ihrer Meinung nach einige aufschlussreiche Einsichten ermögliche:

(T) $(\neg P_c \ \& \ M_c) > B_c$

Laut Le Pore und Loewer kann gezeigt werden, dass (T) sowohl mit dem Anomalen Monismus als auch mit (S) verträglich ist. Zudem sei es sehr wahrscheinlich, dass (T) de facto wahr ist. Gemäß (T) hätte ein Wirkungsereignis auch dann die Eigenschaft B, wenn das Ursacheereignis die Eigenschaft M besessen, gleichzeitig aber nicht P instanziiert hätte. Das Ereignis, mit der Eigenschaft mein Begehren Wasser zu trinken zu sein, hätte nämlich selbst dann, wenn es nicht P gehabt hätte, eine andere physikalische Eigenschaft bzw. Realisierung P* besessen. In der nächsten möglichen Welt, in der ich das Verlangen verspüre meinen Durst zu löschen, dieser Wunsch aber nicht P ist, wird dieses Begehren dennoch mein Armstrecken verursachen. Eventuell gibt es sogar ein Gesetz (wenngleich auch sicherlich kein striktes), das jemand, der ein Verlangen verspürt seinen Durst zu stillen und zugleich glaubt, dass vor ihm ein Glas Wasser steht, das er durch ein Strecken seines Armes erreichen kann, ceteris paribus seinen Arm ausstrecken wird. In dem Fall, in dem ein beliebiges c M hat, zugleich jedoch nicht P besitzt, wäre es eben jenes ceteris paribus Gesetz, das an dieser Stelle einspringt und dafür sorgt, dass c ein Ereignis verursacht, das die Eigenschaft B besitzt. Nach Le Pore und Loewer lassen sich zahllose Beispiele dieser Art auch im nicht-psychologischen Bereich auflisten. Angenommen, folgendes Ereignis tritt ein: ein Hurrikan namens Donald trifft auf die Küste und verursacht, dass Straßen überschwemmt werden. Dieses Ereignis ist identisch mit einem Ereignis, bei dem gewisse Luft- und Wassermoleküle sich auf eine bestimmte, sehr komplexe Art und Weise, bewegen. Die Eigenschaft, aus Molekülen, die sich auf exakt diese Art bewegen, zu bestehen, könnte man beispielsweise als P bezeichnen. Nun ist es durchaus vorstellbar, dass folgendes kontrafaktische Konditional wahr ist: wenn der Hurrikan Donald nicht die Eigenschaft P gehabt hätte, dann wären die Straßen trotzdem überflutet worden. Es ist z.B. vorstellbar, dass dieser Hurrikan dem ursprünglichen Wirbelsturm Donald so ähnlich wie möglich gewesen wäre, aber gerade die Eigenschaft P nicht gehabt hätte. Stattdessen hät-

te besagter Hurrikan etwa die Eigenschaft P^* instanziiert, die der Eigenschaft P hinreichend ähnlich ist und von der gilt, dass P^* -Ereignisse unter geeigneten Bedingungen Überschwemmungen verursachen. Folgt man diesem Gedankengang, dann wäre allerdings Donalds Eigenschaft ein Hurrikan zu sein kausal irrelevant dafür, dass er die Straßen überflutet hat. Dies wiederum widerspricht jeglicher Intuition. Aus solchen Beispielen ziehen Le Pore und Loewer den Schluss, dass (III) eine zu starke Bedingung für kausale Relevanz₂ ist.¹⁴⁹ Natürlich stellt sich sofort die Frage, was ein angemessener Ansatz kausaler Relevanz₂ Le Pore und Loewer zufolge leisten muss. Ihre Antwort fällt folgendermaßen aus: „A fully adequate account of causal relevance₂ should show how mentalistic counterfactuals are grounded.“¹⁵⁰ Die Tauglichkeit einer Definition kausaler Relevanz₂ lasse sich laut Le Pore und Loewer schlichtweg daran bemessen, inwiefern es möglich ist mittels des entsprechenden Ansatzes kontrafaktische Bedingungen, die mentalistisches Vokabular enthalten, vernünftig begründen zu können. Eine Möglichkeit, die in diesem Zusammenhang im Anomalen Monismus gegeben ist, sind nicht-strikte psychophysische und psychologische Gesetze. Die Existenz solcher *ceteris paribus* Gesetze ist schließlich ausdrücklich mit dem Anomalen Monismus verträglich. Schlussendlich können *ceteris paribus* Gesetze kontrafaktische Bedingungen bestätigen. Zur Untermauerung dieser These führen Le Pore und Loewer folgendes Beispiel an:

Let R , W , and B be the statements that a red block, a white block, and a blue block is placed in front of Donald and S_r , S_w , and S_b be the statements that Donald sees a red block, a white block, and a blue block. We will suppose, as is plausible, that there are nonstrict laws of the form:

(L) If X and C , then S_x ,

where C are conditions like lighting is good, Donald is awake and paying attention, and so on. Even with such conditions added, the law is a *ceteris paribus* one and, if AM is correct, it will be impossible to add explicit conditions that turn it into a strict law. When the laws (L) hold, we will say that the statements describing what Donald sees *depend nomologically* on the statements describing the blocks in front of him. Call conditions C *counterfactually independent* of the family of statements $\{R, W, B\}$, if C would continue to hold no matter which member of $\{R, W, B\}$ is true. Lewis shows that, if C and the *ceteris paribus* conditions associated with (L) are counterfactually independent of $\{R, W, B\}$, then S_x will depend counterfactually on X . That is, each of the counterfactuals, $R > S_r$, $W > S_w$, $B > S_b$, will be true. If we further assume that a block which has one of three colors will be placed in front of Donald [and that this statement is also counterfactually independent of $\{R, W, B\}$], then the statement $\neg X > \neg S_x$ will also be true.¹⁵¹

Angenommen R , W und B sind die Behauptungen, dass ein roter, ein weißer bzw. ein blauer Klotz vor Donald liegt und S_r , S_w und S_b sind die Aussagen, dass Donald einen roten, einen weißen oder einen blauen Klotz sieht. Außerdem existiere ein nicht-striktes Gesetz (L) der Form: Wenn X und C , dann S_x , wobei C sämtliche Bedingungen wie z.B. die Lichtverhältnisse sind angemessen, Donald ist wach und aufmerksam, usw. umfasst. Offensichtlich bleibt (L)

¹⁴⁹ Vgl. Le Pore, Ernest & Loewer, Barry: „Mind Matters“ in: *The Journal of Philosophy*, Vol. 84, 1987, S. 639f.

¹⁵⁰ Le Pore, Ernest & Loewer, Barry: „Mind Matters“ in: *The Journal of Philosophy*, Vol. 84, 1987, S. 640.

¹⁵¹ Le Pore, Ernest & Loewer, Barry: „Mind Matters“ in: *The Journal of Philosophy*, Vol. 84, 1987, S. 640f.

aber trotz dieser Zusatzannahmen ein *ceteris paribus* Gesetz und dem Anomalen Monismus zufolge ist es prinzipiell nicht möglich (L) durch das Hinzufügen auch noch so detaillierter Zusatzannahmen in ein striktes Gesetz umzuwandeln. Wenn (L) gültig ist, dann sind die Aussagen, die beschreiben was Donald sieht, nomologisch abhängig von den Aussagen, die die Klötze vor ihm beschreiben. Die Bedingungen C sind kontrafaktisch unabhängig von der Klasse der Aussagen {R, W, B}, falls die Konditionen C gültig sind und zwar unabhängig davon welche der Aussagen R, W, oder B eintritt. In seinem äußerst einflussreichen Aufsatz *Causation* hat David Lewis gezeigt, dass, wenn C und die *ceteris paribus* Annahmen von (L) kontrafaktisch unabhängig von {R, W, B} sind, S_x kontrafaktisch von X abhängt. Geht man zudem davon aus, dass ein Klotz in einer der drei Farben vor Donald gelegt wird, so ist auch die Behauptung $\neg X > \neg S_x$ wahr. Wird beispielsweise ein roter Klotz vor Donald platziert, dann wird schließlich dieses Ereignis das Ereignis ‚Donald sieht einen roten Klotz‘ verursachen. Wenn jedoch das erste Ereignis nicht das Hinlegen eines roten Klotzes gewesen wäre, dann wäre auch das zweite Ereignis nicht Donalds Erblicken eines roten Klotzes gewesen. Dies alles beweist laut Le Pore und Loewer, dass *ceteris paribus* Gesetze in der Lage sind kontrafaktische Aussagen zu bestätigen. Dementsprechend sollte auch die Agenda einer Psychologie, die sich dem anomalen Monismus verpflichtet sieht, in „(...) the discovery and the systematization of such nonstrict laws (at various levels) connecting psychological and/or behavioral properties (...)“¹⁵² bestehen.

Abschließend resümieren Le Pore und Loewer nochmals die zentralen Inhalte ihres Essays und bilanzieren ihren Beitrag zur Verteidigung des Anomalen Monismus:

We have seen that AM attempts to resolve the mind-body problem by endorsing (2), (3), and (3'), denying (1) with respect to events, and affirming (1) with respect to properties. Davidson is silent on (2) and (3) in respect to properties, leading to the accusation that AM is committed to epiphenomenalism. We rebutted this charge by showing that AM is compatible with there being counterfactual dependencies between events in virtue of their mental properties. To do this is to affirm (2) with respect to properties but, of course, to deny (3) with respect to properties. An event's physical features may counterfactually depend on another event's mental features. But, interestingly, we need not deny (3') for our account of causal relevance₂. It may be that all counterfactuals *supervene* on basic physical truths and strict laws. That is, if two possible worlds are exactly alike with respect to basic physical facts and strict laws, they are exactly alike with respect to counterfactuals. This fairly strong physicalism still allows sufficient autonomy of the mental so that it is not reducible to the physical and it has a genuine explanatory and causal role to play.¹⁵³

Um Davidsons Lösung für das Problem um die Anomalie des Mentalen zu untersuchen, ist es hilfreich zunächst nochmals kurz das Trilemma vorzustellen, aus dem sich dieses Problem ableitet. Dabei erscheint jede der folgenden Thesen für sich genommen jeweils als sehr plau-

¹⁵² Le Pore, Ernest & Loewer, Barry: „*Mind Matters*“ in: *The Journal of Philosophy*, Vol. 84, 1987, S. 641.

¹⁵³ Le Pore, Ernest & Loewer, Barry: „*Mind Matters*“ in: *The Journal of Philosophy*, Vol. 84, 1987, S. 641f.

sibel, allerdings hat es den Anschein als sei es nicht möglich alle gemeinsam konsistent zu vertreten:

- (1) Das Mentale und das Physikalische sind unterschiedlich.
- (2) Zwischen Mentalem und Physikalischem findet kausale Interaktion statt.
- (3) Das Physikalische ist kausal geschlossen.
- (3') Jede Art von Verursachung kann auf physikalische Verursachung reduziert werden oder superveniert auf physikalischer Verursachung.

Davidsons Schlüssel zur Lösung dieses Problems besteht im Wesentlichen darin, an (2) und (3) festzuhalten und (1) im Hinblick auf Ereignisse als falsch, im Bezug auf Eigenschaften jedoch als wahr einzustufen. Bei den Thesen (2) und (3) äußert sich Davidson nicht über Eigenschaften. Dies hat ihm den Vorwurf eingebracht, der Anomale Monismus sei im Grunde eine epiphänomenalistische Position. Diesen Einwand weisen Le Pore und Loewer zurück, indem sie zeigen, dass es mit dem Anomalen Monismus vereinbar ist, dass kontrafaktische Abhängigkeiten zwischen Ereignissen aufgrund deren mentaler Eigenschaften bestehen. D.h. Le Pore und Loewer bekräftigen (2) und leugnen (3) im Hinblick auf Eigenschaften. Denn die physikalischen Eigenschaften eines Ereignisses können von den mentalen Eigenschaften eines anderen Ereignisses kontrafaktisch abhängig sein. Außerdem sind Le Pore und Loewer dank ihrer Definition von kausaler Relevanz₂ nicht gezwungen auch (3') abzulehnen. Schließlich ist es durchaus möglich, dass alle kontrafaktischen Bedingungen auf basalen physikalischen Wahrheiten und strikten Gesetzen beruhen. Dies bedeutet, dass zwei mögliche Welten, die sich im Hinblick auf ihre basalen physikalischen Wahrheiten und ihre strikten Gesetze genau gleichen, auch identisch im Bezug auf ihre kontrafaktischen Konditionale sind. Dies führt einerseits zu einem starken Physikalismus, ermöglicht dem Mentalen aber andererseits auch genügend Eigenständigkeit, um nicht vollständig auf das Physikalische reduzierbar zu sein, sowie seiner kausalen und explanatorischen Rolle gerecht zu werden.

Sind Kausalgesetze notwendigerweise strikte Gesetze?

Ein bekannter Kritiker des Problems um die Anomalie des Mentalen ist Jerry Fodor. Er kritisiert, dass die Problematik um den anomalen Charakter des Mentalen im Grunde auf einer falschen Prämisse beruht. Da Fodor sowohl die kausale Wechselwirkung zwischen Mentalem und Physikalischem als auch die Anomalie des Mentalen als Tatsachen betrachtet, richtet sich

seine Kritik am Argument hinsichtlich des Anomalie-Problems auf die zweite Prämisse – nämlich die, dass Ereignisse, die in einem Kausalverhältnis zueinander stehen, ein striktes Gesetz instanziiieren müssen.

Als Ausgangspunkt beklagt Fodor in seinem Aufsatz *Making Mind Matter More* innerhalb der zeitgenössischen Philosophie des Geistes ein Grassieren von sogenannter ‚Epiphobie‘. Unter diesem Terminus versteht er die Furcht sich in einen Epiphänomenalisten zu verwandeln bzw. sich zum Annehmen dieser Sichtweise genötigt zu sehen. Dementsprechend besteht die Hauptsorge von Philosophen, die von der Epiphobie ‚infiziert‘ sind, in der Annahme es sei nicht mit dem Physikalismus vereinbar, dass intentionale Zustände qua Intentionalem für Verhalten kausal verantwortlich sind – vielmehr laufen dieser Hypothese zufolge die tatsächlichen Kausalprozesse ausschließlich auf der physikalischen Ebene ab und ergo sind intentionale Zustände reine Epiphänomene.¹⁵⁴ Allerdings ist die Epiphobie laut Fodor „(...) a neurotic worry; if there is a problem, it is engendered (...) by two philosophical mistakes: (a) a wrong idea about what it is for a property to be causally responsible; and (b) a complex of wrong ideas about the relations between special-science laws and the events that they subsume.“¹⁵⁵ Die Wurzel der Epiphobie verortet Fodor also in zwei weit verbreiteten philosophischen Fehlern: Erstens in einer falschen Vorstellung darüber, was es für eine Eigenschaft bedeutet kausal verantwortlich für etwas zu sein und zweitens in einem Konglomerat von falschen Gedanken hinsichtlich der Beziehung von Gesetzen in Einzelwissenschaften und den Ereignissen, die sie subsumieren.

Als Agenda für seinen Essay legt Fodor folgendes Vorgehen fest: zunächst versucht er aufzuzeigen, warum die Epiphobie einigen Philosophen als vernünftige Position erscheint. Anschließend entwickelt er eine skizzenhafte Theorie darüber, was es für eine Eigenschaft bedeutet kausal verantwortlich für etwas zu sein. Mithilfe dieses Ansatzes will Fodor demonstrieren, dass intentionale Eigenschaften kausal verantwortlich sein können, wenn es intentionale kausale Gesetze gibt. Schließlich argumentiert Fodor, dass es gegen die Existenz solcher intentionaler Kausalgesetze – entgegen Davidsons Annahme – keine guten Einwände gebe.

Eine weitverbreitete Art und Weise der Epiphobie zu verfallen, besteht laut Fodor in der Akzeptanz zweier Prämissen und einer Festsetzung. Diese beschreibt er wie folgt:

1. Premise (Supervenience of Causal Powers): The causal powers of an event are entirely determined by its physical properties. Suppose two events are identical in their physical properties; then all causal hypotheticals true of one event are true of the other. If, for example, e1 and e2 are events identical in their physical properties, then all hypotheticals of the form “if e1 occurred in situation S, it would cause...” remain true if “e2” is substituted for “e1” and vice versa.

¹⁵⁴ Zudem wird eine strikte Identität zwischen intentionalen und physikalischen Zuständen geleugnet.

¹⁵⁵ Fodor, Jerry: ‚*Making Mind Matter More*‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 59.

2. Premise (Property Dualism): Intentional properties supervene on physical properties, but no intentional property is identical to any physical property. (A physical property is a property expressible in the vocabulary of physics. Never mind, for now, what the vocabulary of physics is; just assume that it contains no intentional terms.)
3. Stipulation: A property is “causally responsible” iff it affects the causal powers of things that have it. And (also by stipulation) all properties that aren’t causally responsible are epiphenomenal.¹⁵⁶

Die erste Prämisse, die maßgeblich zur Attraktivität der Epiphobie beiträgt, behauptet die Supervenienz kausaler Kräfte. Dieser These zufolge sind die kausalen Kräfte eines Ereignisses vollständig durch dessen physikalische Eigenschaften determiniert. Wenn zwei Ereignisse in allen physikalischen Aspekten identisch sind, dann stimmen auch alle hypothetischen Kausalaussagen beider Ereignisse überein. Nimmt man beispielsweise an, dass die beiden Ereignisse *e*₁ und *e*₂ in allen physikalischen Belangen äquivalent sind, dann muss in jeder Behauptung der Form ‚Wenn sich *e*₁ in Situation *S* ereignet, würde es ... verursachen‘ ‚*e*₁‘ salva veritate durch ‚*e*₂‘ ersetzt werden können. Die zweite Prämisse, die laut Fodor zum Nimbus der Epiphobie in der zeitgenössischen Philosophie des Geistes beiträgt ist die Annahme eines Eigenschaftsdualismus. Diese These besagt im Wesentlichen, dass intentionale Eigenschaften zwar auf physikalischen Eigenschaften supervenieren, letzten Endes aber nicht identisch mit physikalischen Eigenschaften sind. Der Unterschied intentionaler und physikalischer Eigenschaften liegt in der Art ihrer Beschreibung begründet – physikalische Eigenschaften greifen hierbei ausschließlich auf physikalisches Vokabular zurück. Gemeinsam mit diesen beiden Prämissen verleiht, Fodor zufolge, eine weitere Annahme der Epiphobie eine gewisse Anziehungskraft. Gemäß dieser Behauptung ist eine Eigenschaft genau dann kausal verantwortlich, wenn sie die kausalen Kräfte der Entitäten, die diese Eigenschaft besitzen, bestimmt. Aus dieser Aussage lässt sich ableiten, dass alle Eigenschaften die nicht kausal verantwortlich sind als epiphänomenal einzustufen sind. Anhand folgenden Beispiels demonstriert Fodor wie die beiden Prämissen zusammen mit der Annahme zum epiphänomenalen Charakter intentionaler Eigenschaften führen:

(...) [C]onsider the mental event *m* (let’s say, an event which consists of your desiring to lift your arm) which is the cause of the behavioral event *b* (let’s say, an event which consists of you lifting your arm). *m* does, of course, have certain intentional properties. But, according to 2, none of its intentional properties is identical to any of its physical properties. And, according to 1, *m*’s physical properties fully determine its causal powers (including, of course, its power to cause *b*). So, it appears that *m*’s being the cause of your lifting your arm doesn’t depend on its being a desire to lift your arm; *m* would have caused your lifting of your arm even if it hadn’t had its intentional properties, so long as its physical properties were preserved. So it appears that *m*’s intentional properties don’t affect its causal powers. So it appears that *m*’s intentional properties are causally inert. Clearly, this argument iterates to *any* intentional property of the cause of any behavioral effect. So the intentional properties of mental events are epiphenomenal. Epiphobia!¹⁵⁷

¹⁵⁶ Fodor, Jerry: ‚*Making Mind Matter More*‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 60.

¹⁵⁷ Fodor, Jerry: ‚*Making Mind Matter More*‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 60f.

Angenommen, ein mentales Ereignis *m* (beispielsweise mein Verlangen meinen Arm zu heben) verursacht ein Verhaltensereignis *b* (etwa das Ereignis, dass ich meinen Arm hebe), dann hat *m* sicherlich bestimmte intentionale Eigenschaften (z.B. die Eigenschaft ein Verlangen zu sein). Nach Prämisse 2 kann nun aber keine dieser intentionalen Eigenschaften identisch mit einer physikalischen Eigenschaft von *m* sein. Da aber Prämisse 1 zufolge allein *m*'s physikalische Eigenschaften dessen kausale Kräfte vollständig festlegen, ist somit auch *m*'s kausale Kraft *b* zu verursachen ausschließlich durch seine physikalischen Eigenschaften bestimmt. Daher scheint es als würde *m* als Ursache von meinem Armheben nicht davon abhängen, dass *m* ein Verlangen war meinen Arm zu heben. Schließlich hätte *m* auch dann noch mein Armheben verursacht, wenn es nicht die (oder möglicherweise auch gar keine) intentionalen Eigenschaften gehabt hätte, die es tatsächlich hatte, solange *m* nur genau die physikalischen Eigenschaften instanziiert hat, die es *de facto* instanziiert hat. Somit hat es den Anschein, als hätten *m*'s intentionale Eigenschaften keinerlei Auswirkung auf die in *m* innewohnenden kausalen Kräfte und würden kausal sozusagen ins Leere laufen. Es ist offensichtlich, so Fodor, dass sich dieses Argument auf *jede beliebige* intentionale Eigenschaft, die ein Verhalten verursacht, übertragen lässt. Folglich besitzen diesem Gedankengang zufolge *alle* intentionalen Eigenschaften lediglich einen epiphänomenalen Charakter – et voilà: dies ist nichts anderes als die Kernaussage der Epiphobie.

Allerdings beschränkt sich dieses Argument nicht allein auf intentionale Eigenschaften, sondern lässt sich auf *alle nicht-physikalischen* Eigenschaften ausweiten, solange man von der Gültigkeit des Eigenschaftsdualismus ausgeht. Angenommen, die Eigenschaft ein Berg zu sein ist keine physikalische Eigenschaft.¹⁵⁸ Die ungeschulte Intuition würde es nun zunächst eben dieser Eigenschaft ein Berg zu sein zuschreiben, dass z.B. der Gipfel des Mount Everest schnee- bzw. gletscherbedeckt ist oder dass so viele Menschen versuchen den Mount Everest zu besteigen. Dem soeben vorgestellten Epiphobieargument zufolge, ist dies jedoch nicht der Fall. Denn die kausalen Kräfte des Mount Everest sind vollständig durch dessen physikalische Eigenschaften fixiert. Da allerdings die Eigenschaft ein Berg zu sein keine physikalische Eigenschaft ist, besitzt sie auch keinerlei kausale Kraft. Laut Fodor gibt es Philosophen, die bereit sind diese bittere Pille zu schlucken. Grundsätzlich verfolgen sie dabei zwei Strategien eine solch radikale Behauptung zu rechtfertigen: „Such folk may (i) either deny that property dualism applies to mountainhood (because, on reflection, *being a mountain* is a physical property after all) or (ii) assert that it is intuitively plausible that *being a mountain* is causally inert (because, on reflection, it is intuitively plausible that it's not *being a mountain* but some

¹⁵⁸ D.h. lediglich, dass der Begriff ‚Berg‘ im physikalischen Vokabular nicht vorkommt und dass die Eigenschaft ein Berg zu sein letztendlich nicht vollständig auf eine physikalische Eigenschaft reduzierbar ist.

other of Mt. Everest's properties – specifically, some of its physical properties – that are causally responsible for its effects).¹⁵⁹ Zum Einen könnte man also leugnen, dass der Eigenschaftsdualismus auf die Eigenschaft des Bergseins Anwendung findet, weil die Eigenschaft ein Berg zu sein schlussendlich doch eine physikalische Eigenschaft ist. Zum Anderen könnte man argumentieren, dass es sehr wohl intuitiv plausibel ist die Eigenschaft ein Berg zu sein für kausal wirkungslos zu erachten. Schließlich sollte sich bei näherer Betrachtung herausstellen, dass nicht die Eigenschaft ein Berg zu sein kausal verantwortlich für eine potentielle Wirkung des Mount Everest ist, sondern vielmehr irgendeine seiner physikalischen Eigenschaften. Dem hält Fodor Folgendes entgegen: „(...) [L]et me emphasize that there are lots and lots and *lots* of examples where, on the one hand, considerations like multiple realizability make it implausible that a certain property is expressible in physical vocabulary; and, on the other hand, claims for the causal inertness of the property appear to be widely implausible, at least *prima facie*.“¹⁶⁰ Fodor moniert, es gebe eine überwältigende Menge an Beispielen, in denen es z.B. Überlegungen hinsichtlich multipler Realisierbarkeit, mehr als fragwürdig erscheinen lassen, dass gewisse Eigenschaften sich im Vokabular der Physik ausdrücken lassen. Zugleich wäre es aber auch – zumindest auf den ersten Blick – sehr unglaublich diesen Eigenschaften jegliche kausale Kraft abzusprechen. Zur Illustration führt Fodor folgendes Beispiel an:

Consider the property of being a sail. (...) [S]ails are *airfoils* and there is quite a nice little theory about the causal properties of airfoils. Typically, airfoils generate lift in a direction, and in amounts, that are determined by their geometry, their rigidity, and many, many details of their relations to the (liquid or gaseous) medium through which they move. The basic idea is that the lift is propagated at right angles to the surface of the airfoil along which the medium flows fastest, and is proportional to the relative velocity of the flow. (...)

Now, the relative velocity of the airfoil may be increased by forcing the medium through a “slot” (a constriction, one side of which is formed by the surface of the airfoil). The controlling law is that the narrower the slot, the faster the flow. (...) Anyhow, airfoils and slots can be made out of all sorts of things; sails are airfoils, but so are keel-wings, and airplane wings and bird's wings. Slots are multiply realizable too: You can have a slot both sides of which are made of sailcloth (...), but you can also have a slot one side of which is made of sailcloth and the other side of which is made of *air*. (...) So then, if one of your reasons for doubting that *believing that P* is a physical property is that believing is multiply realizable, then you have the same reason for doubting that *being an airfoil* or *being a slot* counts as a physical property.

And yet, of course, it would seem to be quite mad to say that *being an airfoil* is causally inert. Airplanes fall down when you take their wings off; and sailboats come to a stop when you take down their sails. Everybody who isn't a philosopher agrees that these and other facts are explained by the story about lift being generated by causal interactions between the airfoil and the medium. If that *isn't* the right explanation, what keeps the plane up? If that *is* the right explanation, how could it be that *being an airfoil* is causally inert?¹⁶¹

¹⁵⁹ Fodor, Jerry: ‚*Making Mind Matter More*‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 61.

¹⁶⁰ Fodor, Jerry: ‚*Making Mind Matter More*‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 61.

¹⁶¹ Fodor, Jerry: ‚*Making Mind Matter More*‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 61f.

Zur Untermauerung seiner These wählt Fodor die zunächst gänzlich ‚unschuldig‘ wirkende Eigenschaft ein Segel zu sein. Segel funktionieren wie Tragflächen¹⁶² und es gibt eine Theorie über die kausalen Eigenschaften von Tragflächen. Normalerweise generieren Tragflächen Schub oder besser gesagt dynamischen Auftrieb in eine Richtung und in einer Größenordnung, die von der Form und Starrheit der Tragfläche, sowie von vielen anderen Eigenschaften und deren Beziehungen zu dem Medium, in dem sie sich bewegt, abhängen. Der Grundgedanke ist, dass durch das an der Tragfläche entlang strömende Medium ein Schub bzw. Auftrieb erzeugt wird, der rechtwinklig zur Tragfläche wirkt und proportional zur relativen Geschwindigkeit der Strömung ist. Die relative Geschwindigkeit kann noch dadurch gesteigert werden, dass man das Medium sozusagen dazu ‚zwingt‘ durch eine so genannte ‚Düse‘ zu strömen. Eine solche Düse ist eine Verengung, von der eine Seite durch das Segel geformt wird. Dabei gilt je enger die Düse, desto höher die Strömungsgeschwindigkeit. Nun können Tragflächen und Düsen jedoch auf viele verschiedene Arten realisiert sein – so können beispielsweise Segel, aber auch die Tragflächen eines Flugzeugs oder die Flügel eines Vogels als Tragflächen im obigen Sinne fungieren. Ganz analog können auch Düsen aus ganz unterschiedlichen Bestandteilen zusammengesetzt sein. Eine Düse kann beispielsweise auf beiden Seiten aus Segeltuch bestehen oder aber auf der einen Seite aus Segeltuch und auf der anderen Seite aus Luft.

Wer die Eigenschaft „zu glauben, dass P“ aufgrund ihrer multiplen Realisierbarkeit nicht als physikalische Eigenschaft ansieht, wäre aus demselben Grund gezwungen die Eigenschaft „ein Segel zu sein“ ebenfalls nicht als physikalische Eigenschaft zu betrachten. Allerdings wäre es selbst dann – also falls man der Eigenschaft „ein Segel zu sein“ den Status einer physikalischen Eigenschaft abspricht – abwegig diese Eigenschaft zur gleichen Zeit auch als kausal wirkungslos zu erachten. Schließlich würden Flugzeuge von Himmel fallen, wenn man die Tragflächen während des Fluges abmontieren würde und Segelboote würden stehenbleiben sollte man deren Segel einholen. In diesem Zusammenhang würde wohl jedermann – abgesehen von einigen hartgesottenen Philosophen – das Abstürzen der Flugzeuge und das Stehenbleiben der Schiffe mit dem Ausbleiben des dynamischen Auftriebs, der zuvor von den kausalen Wechselwirkungen zwischen Tragfläche und Luft erzeugt wurde, erklären. Daher steht man vor folgendem Dilemma: Liefert die Eigenschaft eine Tragfläche zu sein keine gute Erklärung für das Fliegen des Flugzeugs, dann stellt sich sofort die Frage welcher anderen Eigenschaft es sonst zu verdanken wäre, dass das Flugzeug in der Luft verweilt und nicht abstürzt. Gibt die Eigenschaft eine Tragfläche zu sein hingegen eine gute Erklärung dafür an die

¹⁶² Dies gilt nicht für Rahsegel, sondern nur für Schratsegel und insbesondere dann, wenn sie in einem Anstellwinkel von etwa 5° bis 30° angeströmt werden.

Hand, dann ist vollkommen unklar, warum bzw. wie diese Eigenschaft zugleich auch kausal wirkungslos sein könnte.

Allerdings macht das Epiphobieargument nicht bei intentionalen Eigenschaften halt, sondern lässt sich ohne Modifikation auf alle Eigenschaften übertragen, die sich nicht im physikalischen Vokabular ausdrücken lassen und die gleichzeitig nicht identisch mit einer physikalischen Eigenschaft sind. Dies hat zur Folge, dass zumindest ein Großteil der Eigenschaften, die in den Spezialwissenschaften eine kausale Rolle innehaben, nach diesem Verständnis faktisch nur mit einem epiphänomenalen Charakter ausgestattet ist. Deshalb konstatiert Fodor: „Surely something *must* have gone wrong with arguments that show that all these properties are epiphenomenal. How could there be laws about airfoils (notice, laws about the *causal consequences of something's being an airfoil*) if airfoilhood is epiphenomenal?“¹⁶³ Ein Argument, dessen Konklusion nahezu alle Eigenschaften als epiphänomenal ausweist muss nach Fodor einen essentiellen Fehler beinhalten. Wie könnte es ansonsten Gesetze über die kausalen Implikationen der Eigenschaft eine Tragfläche zu sein geben, wenn diese Eigenschaft epiphänomenal ist? Den Schlüssel gegen das Epiphobieargument sieht Fodor in einem adäquaten Verständnis darüber, was es für eine Eigenschaft heißt kausal verantwortlich zu sein. Denn: „(...) it ought to be a minimal condition upon a theory of what it is for something to be a causally responsible property that it does not entail the epiphenomenality of winghood, mountainhood, gearhood, leverhood, beliefhood, desirehood and the like.“¹⁶⁴ Schließlich scheint es vernünftig als Mindestvoraussetzung für eine Theorie über die kausale Verantwortung von Eigenschaften festzulegen, dass die Eigenschaften eine Tragfläche, ein Berg, ein Glaube, ein Verlangen, etc. zu sein nicht epiphänomenal sein dürfen. Sollte es gelingen einen solchen Ansatz von kausaler Verantwortung zu formulieren, wäre das nicht nur ein Tonikum gegen die Epiphobie hinsichtlich intentionaler Eigenschaften, sondern gegen die Epiphobie tout court. Dabei genüge es laut Fodor zu zeigen, dass intentionale Eigenschaften die notwendigen Bedingungen für kausale Verantwortung erfüllen können – ein Nachweis darüber, dass sie zugleich auch den hinreichenden Bedingungen kausaler Verantwortung gerecht werden, sei nicht nötig. Außerdem gelte es die Frage ‚Was macht eine Eigenschaft kausal verantwortlich?‘ von der deutlich schwierigeren Frage ‚Was bestimmt welche Eigenschaft kausal verantwortlich ist, wenn ein Ereignis ein anderes verursacht?‘ zu unterscheiden. Wenn ein Ereignis e1 ein Ereignis e2 verursacht, so tut es dies aufgrund der einen oder anderen Eigenschaft, die kausal verantwortlich ist (d.h. kraft derer e1 eine Ursache sein kann). Möglicherweise hat e1 aber nicht nur eine, sondern zahlreiche solcher Eigenschaften. Daher ist es ein Trugschluss

¹⁶³ Fodor, Jerry: ‚*Making Mind Matter More*‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 63.

¹⁶⁴ Fodor, Jerry: ‚*Making Mind Matter More*‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 63.

anzunehmen, dass, wenn es e1 wegen einer bestimmten Eigenschaft P möglich ist eine Ursache zu sein, es allein aufgrund dieser Tatsache auch die Eigenschaft P sein muss, vermittels derer e1 e2 verursacht. Tatsächlich würde selbst wenn es e1 aufgrund der Eigenschaft P möglich wäre die Ursache von e2 zu sein, nicht zwangsläufig daraus folgen, dass e1 e2 verursacht hat, weil e1 P instanziierte. Schließlich könnte es weiterhin zahlreiche andere Eigenschaften geben, die es e1 auch ermöglichen die Ursache von e2 zu sein und dabei muss es nicht offenkundig sein aufgrund welcher dieser Eigenschaften e1 in praxi die Ursache von e2 ist. Dementsprechend unterscheidet Fodor grundsätzlich zwei Möglichkeiten wie mittels intentionaler Eigenschaften Verhalten erklärt werden könne. Zum Einen könnte man zeigen, dass mein Verlangens meinen Arm zu heben kraft seines intentionalen Gehalts etwas in der Welt (nämlich das Heben meines Arms) bewirkt hat. Zum Anderen könnte nachgewiesen werden, dass die Eigenschaft ein Verlangen zu sein meinen Arm zu heben die Art von Eigenschaft sein kann, vermittels derer ein Geschehen bewirkt werden könnte. Offensichtlich impliziert die zweite Variante eine deutlich schwächere Aussage – laut Fodor reiche diese Art von Nachweis aber vollkommen aus, um die Epiphobie erfolgreich zu „kurieren“.¹⁶⁵

Anschließend gibt Fodor die Voraussetzungen an, auf denen seine Theorie kausaler Verantwortung fußt:

ASSUMPTIONS I assume that singular causal statements need to be covered by causal laws. That means something like:

- 4. Covering Principle: If an event e1 causes an event e2, then there are properties F, G such that;
 - 4.1. e1 instantiates F;
 - 4.2. e2 instantiates G;
 - and
 - 4.3. "F instantiations are sufficient for G instantiations" is a causal law.

When a pair of events bears this relation to a law, I'll say that the individuals are each *covered* or *subsumed* by that law and I'll say that the law *projects* the properties in virtue of which the individuals are subsumed by it.¹⁶⁶

Eine wesentliche Grundannahme Fodors besteht also darin, dass singuläre Kausalaussagen stets von einem Kausalgesetz abgedeckt sein müssen. Das Deckungsprinzip besagt daher, dass es immer wenn ein Ereignis e1 ein Ereignis e2 verursacht zwei Eigenschaften F und G geben muss, so dass gilt: e1 instanziiert F, e2 instanziiert G und es existiert ein Kausalgesetz der Form ‚Instanziiierungen von F sind hinreichend für Instanziiierungen von G‘. Wenn zwei Ereignisse in einer solchen Art von Beziehung zu einem Gesetz stehen, dann werden die beiden Einzelereignisse von dem Gesetz gedeckt oder können in anderen Worten unter dieses Gesetz subsummiert werden. Eine weitere Redeweise, die Fodor verwendet um diesen Sach-

¹⁶⁵ Vgl. Fodor, Jerry: ‚*Making Mind Matter More*‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 63f.

¹⁶⁶ Fodor, Jerry: ‚*Making Mind Matter More*‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 64.

verhalt auszudrücken ist: das Gesetz projiziert die Eigenschaften kraft derer die Einzelereignisse unter das Gesetz subsummiert werden können (bzw. bildet diese ab).

Nachdem Fodor die Präliminarien abgeklärt hat, präsentiert er seine Vorstellung hinsichtlich der hinreichenden Bedingungen für die kausale Verantwortung von Eigenschaften:

(...) I can now tell you my sufficient condition for a property to be causally responsible:

5. P is a causally responsible property if it's a property in virtue of which individuals are subsumed by causal laws; or equivalently,
 - 5.1. P is a causally responsible property if it's a property projected by a causal law; or equivalently (since the satisfaction of the antecedent of a law is ipso facto nomologically sufficient for the satisfaction of its consequent),
 - 5.2. P is a causally responsible property if it's a property in virtue of the instantiation of which the occurrence of one event is nomologically sufficient for the occurrence of another;

If this is right, then intentional properties are causally responsible in case there are intentional causal law; aerodynamic properties are causally responsible in case there are aerodynamic causal laws (...). To all intents and purposes, on this view the question whether the property P is causally responsible *reduces to* the question whether there are causal laws about P. To settle the second question *is* to settle the first.¹⁶⁷

Eine Eigenschaft P ist laut Fodor genau dann kausal verantwortlich, wenn P entweder eine Eigenschaft ist aufgrund derer ein Einzelereignis unter ein Kausalgesetz subsummiert werden kann, oder, wenn P eine Eigenschaft ist, die von einem Kausalgesetz projiziert wird, oder wenn P schließlich eine Eigenschaft ist, deren Instanziierung das Eintreten eines Ereignisses nomologisch notwendig für das Eintreten eines anderen Ereignisses macht. Daraus folgt, dass intentionale Eigenschaften genau dann kausal verantwortlich sind, wenn es intentionale Kausalgesetze gibt (dementsprechend sind aerodynamische Eigenschaften genau dann kausal verantwortlich, wenn es aerodynamische Kausalgesetze gibt, usw.). Nach dieser Auffassung lässt sich die Frage nach der kausalen Verantwortung einer Eigenschaft auf die Frage zurückführen, ob Kausalgesetze im Hinblick auf diese Eigenschaft existieren. Mit einer Antwort auf die zweite Frage ist zugleich auch eine Antwort auf die erste gegeben. Den Vorwurf dieser Ansatz sei zirkulär, weil er *kausale* Verantwortung von Eigenschaften mithilfe eines *Kausalgesetzes* erkläre, das die Instanziierung der Eigenschaft abdeckt, verwehrt sich Fodor. Im Sinne dieses Kritikpunktes könnte es beispielsweise auch durchaus möglich sein, dass das Gesetz, das den Fall abdeckt kein Kausalgesetz ist. Ein solches nicht kausales Gesetz ist z.B. in dem Zusammenhang von Druck und Volumen bei einem Gas zu sehen: Druck und Volumen in einem Gas verhalten sich umgekehrt proportional zueinander – wächst das Volumen eines Gases, so schrumpft dessen Druck (bei ansonsten gleichbleibenden Bedingungen) und umgekehrt. Nun könnte man weiter behaupten, dass ein Ereignis, das unter ein solches (nicht kausales) Gesetz fällt, nicht hinreichend für eine kausal verantwortliche Eigenschaft ist. Dieser

¹⁶⁷ Fodor, Jerry: 'Making Mind Matter More' in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 65.

Argumentation hält Fodor entgegen, dass allem Anschein nach keine Eigenschaft ausschließlich in nicht kausalen Gesetzen eine Rolle spielt. Wenn dem aber so ist, dann sind kausal verantwortliche Eigenschaften vermutlich genau die Art von Eigenschaft, aufgrund derer ein Gesetz ein Einzelereignis abgedeckt.¹⁶⁸

Nach Fodors Verständnis ermöglichen die Kausalgesetze der Spezialwissenschaften ebenso sehr wie diejenigen der grundlegenden Wissenschaft (in anderen Worten der Physik) eine Zuschreibung von kausaler Verantwortung. Daher sind die Eigenschaften der Spezialwissenschaften auch selbst dann nicht epiphänomenal, wenn sie auf basaleren Eigenschaften supervenieren. Allerdings weist Fodor auf einen aufschlussreichen Unterschied zwischen diesen beiden Arten von Kausalgesetzen hin: „(...) a metaphysically interesting difference between basic and nonbasic laws is that, in the case of the latter but not the former, there always has to be *a mechanism in virtue of which* the satisfaction of its antecedent brings about the satisfaction of its consequent. If ‘Fs cause Gs’ is basic, then there is no answer to the question *how* do Fs cause Gs; they just do, and that they do is among the not-to-be-further-explained facts about the way the world is put together. Whereas, if ‘Fs cause Gs’ is *nonbasic*, then there is always a story about what goes on when – and in virtue of which – Fs cause Gs.“¹⁶⁹ Nicht-basale Gesetze sind also im Gegensatz zu fundamentalen Gesetzen dadurch gekennzeichnet, dass stets ein Mechanismus am Werk ist, kraft dessen das Eintreten des Antezedens’ das Eintreten des Konsequens’ bewirkt. Falls das Gesetz ‚Fs verursachen Gs’ also ein basales Gesetz ist, dann gibt es keine Antwort auf die Frage wie die Verursachung von einem G durch ein F genau geschieht. Dass dies der Fall ist hängt vielmehr mit dem So-Sein der aktuellen Welt zusammen. Es ist eben schlichtweg so, dass Fs Gs verursachen und dies kann auch nicht weiter erklärt werden. Ist das Gesetz ‚Fs verursachen Gs’ hingegen nicht basal, dann ist es grundsätzlich möglich das Wie der Verursachungsrelation ‚F verursacht G’ genauer zu explizieren. Anschließend untersucht Fodor den Zusammenhang zwischen Gesetzen, kausal verantwortlichen Eigenschaften und implementierenden Mechanismen genauer:

It is, I should add, not in the least unusual to find that the vocabulary that’s appropriate to articulate a special-science law is systematically different from the vocabulary that’s appropriate to articulate its implementing mechanism(s). Rather, shift of vocabulary as one goes from the law to the mechanism is the *general* case. If you want to talk laws of inheritance, you talk recessive traits and dominant traits and homozygotes and heterozygotes; if you want to talk mechanisms of inheritance, you talk chromosomes and genes and how the DNA folds. If you want to talk psychological law, you talk intentional vocabulary; if you want to talk psychological mechanism, you talk syntactic (or maybe neurological) vocabulary. (...) It doesn’t follow that the property of being a belief or an airfoil or a recessive trait is causally inert; all that follows is that *specifying the causally responsible macroproperty isn’t the same as specifying the implementing micromechanism*.

¹⁶⁸ Vgl. Fodor, Jerry: ‚Making Mind Matter More’ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 65f.

¹⁶⁹ Fodor, Jerry: ‚Making Mind Matter More’ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 66.

It's a confusion to suppose that, if there's a law, then there needn't be an implementing mechanism; and it's a confusion to suppose that, if there's a mechanism that implements a law, then the properties that the law projects must be causally inert. If you take great care to avoid both these confusions, you will be delighted to see how rapidly your epiphobia disappears.¹⁷⁰

Laut Fodor unterscheidet sich der Wortschatz, den man zur Beschreibung der Gesetze einer Spezialwissenschaft verwendet, in der Regel erheblich von dem Vokabular, dessen man sich bei der Darstellung des entsprechenden Implementierungsmechanismus' bedient. Wenn man zum Beispiel die Gesetze der Genetik thematisieren will, dann spricht man über rezessive und dominante Merkmale, sowie über homo- und heterozygote Allele. Kommt die Sprache hingegen auf die Mechanismen der Vererbung, so handelt die Rede nun vielmehr von Chromosomen, Genen und DNS. Diskutiert man über psychologische Gesetze, so findet intentionales Sprachgut Verwendung, werden hingegen die Mechanismen in Augenschein genommen, die psychologische Gesetze realisieren, dann greift man stattdessen auf syntaktisches oder neurologisches Vokabular zurück. Aus diesen unterschiedlichen Beschreibungsformen lässt sich jedoch nicht ableiten, dass beispielsweise die Eigenschaften ein rezessives Merkmal oder eine Annahme zu sein kausal wirkungslos sind. Was aus diesen Beobachtungen folgt ist lediglich, dass das Bestimmen der kausal verantwortlichen (Makro-)Eigenschaft nicht identisch damit ist, den implementierenden Mikromechanismus aufzudecken.

Im Grunde beruht die Epiphobie auf einer fehlerhaften Auffassung über den Zusammenhang von kausal verantwortlichen Eigenschaften, Gesetzen und Mechanismen, die diese Gesetze realisieren, so Fodor. Die Annahme, dass es keinen implementierenden Mechanismus geben müsse solange es ein Gesetz gibt, ist dementsprechend genauso falsch wie diejenige, dass falls ein solcher Mechanismus existiert die Eigenschaften, die das Gesetz projiziert kausal wirkungslos sein müssen.

Nachdem Fodor die grundlegenden Begrifflichkeiten geklärt hat, macht er sich daran das Problem um die Anomalie des Mentalen zu analysieren. Ähnlich wie Davidson beschreibt er das Problem anhand dreier sich gegenseitig vermeintlich ausschließender, für sich genommen jedoch jeweils plausibler, Prinzipien:

There seems to be some tension between the following three principles, each of which I take to be *prima facie* sort of plausible:

6. Strict covering: Just like 4 except with the following in place of 4.3.; "P1 instantiations are causally sufficient for P2 instantiations" is a *strict* causal law.
7. Anomia of the mental: The only strict laws are laws of physics. Specifically, there are no strict 'psychophysical' laws relating types of brain states to types of intentional states; and there are no strict 'psychological' laws relating types of mental events to one another or to types of behavioral outcomes.
8. Causal responsibility of the mental: Intentional properties aren't epiphenomenal.

¹⁷⁰ Fodor, Jerry: 'Making Mind Matter More' in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 67f.

(...) So now we can see where the tension between the three principles (6–8) arises. The responsibility of the mental requires covering by intentional laws. But given the revised notion of covering, according to which causes have to be covered by *strict* laws, it must be *physical* laws, and not intentional ones, that cover mental causes. So it turns out that the intentional properties are causally inert even according to the count of causal responsibility commended in part I. ¹⁷¹

Die drei Prinzipien, in deren Umfeld das Spannungsverhältnis hinsichtlich der Anomalie des Mentalen Fodor zufolge zu Tage tritt sind das strikte Deckungsprinzip, das Prinzip der Anomalie des Mentalen und schließlich das Prinzip der kausalen Verantwortung des Mentalen. Das strikte Deckungsprinzip ist eine Erweiterung des Deckungsprinzips, das Fodor in seinen Vorüberlegungen hinsichtlich kausaler Verantwortung von Eigenschaften expliziert hat. Dabei wird das ursprüngliche Deckungsprinzip nur insofern modifiziert, als dass das Kausalgesetz, das bei einer Verursachungsrelation instanziiert wird, nun strikt sein muss. D.h. kausale Beziehungen müssen von einem Gesetz abgedeckt werden, das ausnahmslos gültig ist: immer wenn das Antezedens eintritt, müssen dadurch nomologisch hinreichende Bedingungen geschaffen werden, die das Eintreten des Konsequens herbeiführen. Das Anomalieprinzip besagt im Wesentlichen, dass sich die Gesetze der Physik in ihrer Beschaffenheit essentiell von den Gesetzen der Spezialwissenschaften unterscheiden: während erstere nämlich ohne Ausnahme Gültigkeit besitzen, so trifft dies auf letztere nicht zu, schließlich beinhalten diese Kausalgesetze stets eine *ceteris paribus* Klausel. In anderen Worten gilt die Kernaussage eines Gesetzes der Physik unter allen Umständen, wohingegen die Essenz eines Gesetzes der Spezialwissenschaften nur in den Fällen, in denen die Rahmenbedingungen identisch sind, ihre Gültigkeit besitzt. Das Prinzip der kausalen Verantwortung mentaler Eigenschaften wiederum kann auf die Bedingung reduziert werden, dass mentale Ursachen von intentionalen Kausalgesetzen abgedeckt werden. An diesem Punkt exemplifiziert sich auch das Spannungsverhältnis zwischen den drei genannten Prinzipien: ist die Instanziierung einer mentalen Eigenschaft eine Ursache, dann muss nach dem Prinzip der kausalen Verantwortung diese Kausalrelation von einem intentionalen Kausalgesetz abgedeckt sein. Die abgewandelte Variante des Deckungsprinzips besagt aber wiederum, dass Kausalgesetze, die Verursachungsbeziehungen abdecken, strikt sein müssen, d.h. sie müssen ohne Ausnahme gelten. Gemäß des Anomalieprinzips können aber einzig und allein physikalische Gesetze strikt sein – so etwas wie strikte psychophysische oder strikte psychologische Gesetze existieren nicht. Wenn jedoch mentale Ursachen von strikten Gesetzen abgedeckt sein müssen und ausschließlich physikalische Gesetze strikt sein können, dann ist es unmöglich, dass intentionale Kausalgesetze mentale Ursachen

¹⁷¹ Fodor, Jerry: ‚*Making Mind Matter More*‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 68f.

abdecken. Aus diesem Grund müssen intentionale Eigenschaften selbst dann kausal wirkungslos sein, wenn Fodors Theorie über die kausale Verantwortung richtig ist.

Als Weg aus dieser Zwangslage heraus, schlägt Fodor vor die drei genannten Prinzipien nochmals eingehender zu examinieren. Das Anomalieprinzip scheint ihm dabei über jeden Zweifel erhaben zu sein – schließlich gelten die Gesetze der Spezialwissenschaften tatsächlich nur unter angemessen idealisierten Bedingungen, bzw. wenn störende Wirkungen interagierender Parameter vernachlässigt werden. Als nächstes begutachtet Fodor das Prinzip der kausalen Verantwortung des Mentalen und bemerkt zu diesem Punkt: „Surely we want 8 to come out true on *some* reasonable construal. I’ve opted for a robust reading: mental properties are causally responsible because they are the properties in virtue of which mental causes are subsumed by covering laws; which is to say that mental properties are causally responsible because there are intentional generalizations which specify nomologically sufficient conditions for behavioral outcomes. But this reading of 8 looks incompatible with 7. 7 suggests that there *aren’t* intentionally specifiable sufficient conditions for behavioral outcomes since, at best, intentional laws hold *ceteris paribus*.“¹⁷² Dieses Zitat erhellt, dass für Fodor eine Ablehnung des Prinzips der kausalen Verantwortung mentaler Eigenschaften schlichtweg keine vernünftige Option ist. Schließlich stünde dabei einfach bei weitem zu viel auf dem Spiel – hätten mentale Eigenschaften lediglich einen epiphänomenalen Charakter, dann käme dies einer gravierenden Erschütterung oder sogar einem vollständigen Kollaps des menschlichen Selbst- und Weltverständnisses in seinen Grundfesten gleich.¹⁷³ Zudem lehnt Fodor auch eine Modifikation seines Verständnisses kausaler Verantwortung ab – so ist ihm zufolge eine Lösung des kausalen Hinreichendseins des Intentionalen für Äußerungen im Verhalten nicht möglich. Stattdessen liege der Weg zur Auflösung des Spannungsverhältnisses zwischen dem strikten Deckungsprinzip, dem Anomalieprinzip und dem Prinzip der kausalen Verantwortung des Mentalen, laut Fodor, vielmehr in einer: „(...) route that doesn’t require the subsumption of causes by strict laws as a lemma.“¹⁷⁴ In anderen Worten rückt Fodor das strikte Deckungsprinzip in den Fokus seiner Aufmerksamkeit. Dabei steht die Frage im Vordergrund, wie es möglich ist:

(...) to square the idea that Ms are nomologically sufficient for Bs with the fact that psychological laws are hedged. How can you have it *both* that special laws only necessitate their consequents *ceteris paribus* *and* that we must get Bs *whenever* we get Ms. Answer: you can’t. But what you can have is just as good: viz., that if it’s a law that $M \rightarrow B$ *ceteris paribus*, then it follows that you get Bs whenever you get Ms *and* the *ceteris paribus* conditions are satisfied. This shows us how *ceteris paribus* laws can do serious scientific business, since it cap-

¹⁷² Fodor, Jerry: ‚*Making Mind Matter More*‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 69f.

¹⁷³ Vgl. Fodor, Jerry: ‚*Making Mind Matter More*‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 77.

¹⁷⁴ Fodor, Jerry: ‚*Making Mind Matter More*‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 72.

tures the difference between the (substantive) claim that Fs cause Gs ceteris paribus, and the (empty) claim that Fs cause Gs except when they don't.¹⁷⁵

Fodor ist also insbesondere an einem Abgleich zwischen den beiden Annahmen, dass einerseits Ms nomologisch hinreichend fürs Bs sein sollen, andererseits aber psychologische Gesetze stets nur eingeschränkt gültig sein dürfen, interessiert. Kann es sein, dass die Gesetze der Spezialwissenschaften durchweg ceteris paribus Klauseln enthalten und trotzdem Bs immer zwangsläufig auf Ms folgen müssen. Fodor negiert dies, kommt aber zu dem Schluss, dass das, was in diesem Zusammenhang tatsächlich möglich ist, nahezu genauso gut ist. Wenn nämlich das Gesetz $M \rightarrow B$ ceteris paribus gültig ist, dann folgt daraus, dass Bs immer zwangsläufig auf Ms folgen, wenn die ceteris paribus-Klausel erfüllt ist. Daraus leitet sich auch der Beitrag ab, den ceteris paribus Gesetze zur Wissenschaft leisten können. Denn an dieser Stelle wird ein elementarer Unterschied zwischen der gehaltvollen Behauptung, dass Fs Gs ceteris paribus verursachen, und der nichtssagenden Aussage, dass Fs Gs verursachen, außer wenn sie dies nicht tun, explizit. Dabei muss man nicht einmal in der Lage sein alle ceteris paribus Bedingungen angeben zu können, so Fodor, es reiche bereits aus einige Fälle zu erkennen, in denen sie erfüllt sind. Aus diesem Grund können die ceteris paribus Bedingungen eingeschränkt gültiger Gesetze, selbst wenn sie nicht vollständig aufgezählt werden können, bisweilen durchaus erfüllt sein. Dieser Punkt werde häufig übersehen und darauf gründe auch die vermeintliche Plausibilität von Davidsons Annahme, dass eingeschränkt gültige Gesetze keine Ursachen abdecken können. Alles in allem fasst Fodor seine Gedanken hinsichtlich Davidsons Argument, das ausgehend von der Anomalie des Mentalen auf die Wahrheit des Anomalen Monismus schließt, folgendermaßen zusammen:

Where does all this leave us with respect to the classical Davidsonian argument that infers physicalism from the anomalousness of the mental? It seems to me that we are now lacking any convincing argument for accepting principle 6. Suppose it's true that causes need to be covered by laws that necessitate their consequents; it doesn't follow that they need to be covered by *strict* laws. Hedged laws necessitate their consequents in worlds where their ceteris paribus conditions are satisfied. Why, then, should mental causes that are covered by hedged intentional laws with satisfied antecedents and satisfied ceteris paribus conditions require *further* covering by a strict law of physics?¹⁷⁶

Davidsons Lösung des Problems um die Anomalie des Mentalen und sogar das Problem selbst, beruhen nach Fodor also auf einer falschen Prämisse – und zwar der Annahme, dass Kausalbeziehungen immer von einem strikten Gesetz abgedeckt sein müssen. Vielmehr genüge es, wenn (mentale) Ursachen von eingeschränkt gültigen (intentionalen) Gesetzen abgedeckt werden, solange zugleich sowohl deren Antezedens als auch deren ceteris paribus Be-

¹⁷⁵ Fodor, Jerry: 'Making Mind Matter More' in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 73.

¹⁷⁶ Fodor, Jerry: 'Making Mind Matter More' in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 74f.

dingungen erfüllt sind. In diesem Zusammenhang ist Fodors Leitgedanke, dass sich eingeschränkt gültige Gesetze, falls ihre *ceteris paribus* Klauseln erfüllt sind, genauso verhalten wie strikte Gesetze. Deshalb ist es auch keine notwendige Voraussetzung für Kausalbeziehungen, dass sie ausschließlich durch strikte Gesetze abgedeckt sein müssen. Ebenso sehr ist es Fodor zufolge nämlich möglich, dass eingeschränkt gültige Gesetze diese Rolle ausfüllen. Das rührt daher, dass: „Strict laws are just the special case of hedged laws where the *ceteris paribus* clauses are discharged *vacuously*: they’re the hedged laws for which ‘all else’ is *always* equal.“¹⁷⁷ Genau genommen besteht nach Fodor also kein qualitativer, sondern lediglich ein quantitativer Unterschied zwischen *ceteris paribus* Gesetzen und strikten Gesetzen. Strikte Gesetze sind sozusagen ein Sonderfall von *ceteris paribus* Gesetzen, bei dem die ‚unter sonst gleichen Bedingungen’-Klausel immer erfüllt ist. Der metaphysisch viel interessantere Unterschied besteht laut Fodor nicht zwischen den Gesetzen der Spezialwissenschaften und den Gesetzen der Physik – also zwischen strikten und nicht-strikten Gesetzen – sondern vielmehr in der Dichotomie zwischen basalen und nicht-basalen Gesetzen. Falls $M \rightarrow P$ ein nicht-basales Gesetz ist, dann gibt es eine Antwort auf die Frage *wie* Instanziierungen von *M* Instanziierungen von *P* herbeiführen – meist geschieht dies über eine Veränderung von Mikrostrukturen. Kurzum bei nicht-basalen Gesetzen ist stets ein Implementierungsmechanismus am Werk, wohingegen dies bei basalen Gesetzen nicht der Fall ist. Über den Zusammenhang zwischen nicht-basalen und nicht-strikten Gesetzen spekuliert Fodor folgendermaßen: „It is therefore surely no accident that *hedged* laws are typically – maybe always – *not* basic. On the one hand, it’s intrinsic to a law being hedged that it is nomologically possible for its *ceteris paribus* conditions not to be satisfied. And, on the other hand, a standard way to account for the failure of a *ceteris paribus* condition is to point to the breakdown of an intervening mechanism.“¹⁷⁸ Nicht-strikte Gesetze sind also nicht zufälligerweise meistens, oder vielleicht sogar immer, nicht-basal – denn falls ein Gesetz nicht-strikt ist, so ist dies gleichbedeutend damit, dass ein Versagen seiner *ceteris paribus* Klausel nomologisch möglich ist. Andererseits besteht eine gängige Erklärung für das Scheitern einer *ceteris paribus* Bedingung im Ausbleiben eines bestimmten Mechanismus’. An dieser Stelle hält Fodor fest, dass erstens nicht-basale Gesetze immer eines intervenierenden Mechanismus bedürfen und dass zweitens keine basalen Gesetze der Psychologie existieren. Nimmt man nun noch die Behauptung hinzu, dass alle Mechanismen, auf denen das Wirken von nicht-basalen Gesetzen schlussendlich beruht, physikalisch sind, dann müssen auch die Mechanismen mentaler Verursachung letzten Endes physikalisch sein. Abschließend kontrastiert Fodor seine Theorie mit der Davidsons im

¹⁷⁷ Fodor, Jerry: ‚*Making Mind Matter More*’ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 75.

¹⁷⁸ Fodor, Jerry: ‚*Making Mind Matter More*’ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 76.

Hinblick auf genau diesen Punkt: „(...) my story gives us both physicalism and a reasonable account of the causal responsibility of the mental; whereas Davidson’s story gives us at most the former. But if we *can’t* get both the causal responsibility of the mental and an argument for physicalism, then it seems to me that we ought to give up the argument for physicalism. I’m not really convinced that it matters very much whether the mental is physical (...).“¹⁷⁹ Während seine Theorie also sowohl ein angemessenes Verständnis der kausalen Verantwortung des Mentalen ermögliche und darüber hinaus noch ein Argument für die Wahrheit des Physikalismus liefere, gelinge Davidsons Theorie bestenfalls nur letzteres. Für den Fall aber, dass beides – eine adäquate Erklärung der kausalen Verantwortlichkeit des Mentalen und ein Argument für den Physikalismus – nicht gleichzeitig zu haben ist, so müsse man laut Fodor unbedingt das Festhalten am Physikalismus aufgeben. Schließlich spiele es wahrscheinlich nur eine untergeordnete Rolle, ob das Mentale im Grunde physikalisch ist oder nicht.

¹⁷⁹ Fodor, Jerry: ‚*Making Mind Matter More*‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989, S. 77.

Das Exklusionsproblem

Das Exklusionsproblem ist ein Problem mentaler Verursachung, das sich insbesondere nicht-reduktiven, physikalistischen Ansätzen¹⁸⁰ stellt. Denn nimmt man an, dass mentale Eigenschaften nicht zur Gänze auf physikalische Eigenschaften reduziert werden können, geht zugleich aber davon aus, dass mentale Eigenschaften durch physikalische Eigenschaften realisiert werden, dann stellt sich die Frage, welche kausale Arbeit für die mentalen Eigenschaften übrig bleibt bzw. ob letztere überhaupt noch eigenständig (d.h. in ihrer Eigenschaft mental zu sein, also qua Mentalem) kausal wirksam sein können.

Auch dieses Problem mentaler Verursachung bearbeitet Jaegwon Kim¹⁸¹ ausführlichst. In seinem Buch *Mind in a Physical World* fasst Kim das Exklusionsproblem wie folgt zusammen:

Suppose (...) that mental event *m*, occurring at time *t*, causes physical event *p*, and let us suppose that this causal relation holds in virtue of the fact that *m* is an event of mental kind *M* and *p* an event of physical kind *P*. Does *p* also have a physical cause at *t*, an event of some physical kind *N*?

To acknowledge mental event *m* (occurring at *t*) as a cause of physical event *p* but deny that *p* has a physical cause at *t* would be a clear violation of the causal closure of the physical domain, a relapse into Cartesian interactionist dualism which mixes physical and nonphysical events in a single causal chain. But to acknowledge that *p* has also a physical cause, *p*^{*}, at *t* is to invite the question: Given that *p* has a physical cause *p*^{*}, what causal work is left for *m* to contribute? The physical cause therefore threatens to exclude, and preempt, the mental cause. This is the problem of causal exclusion. The antireductive physicalist who wants to remain a mental realist, therefore, must give an account of how the mental cause and the physical cause of one and the same event are related to each other. Token physicalism, like Davidson's anomalous monism, is not enough, since the question ultimately involves the causal efficacy of mental properties, and antireductionism precludes their reductive identification with physical properties. Thus the problem of causal exclusion is to answer this question: *Given that every physical event that has a cause has a physical cause, how is a mental cause also possible?*¹⁸²

Das Exklusionsproblem entsteht also in folgendem Zusammenhang: Angenommen ein mentales Ereignis *m* verursacht ein physikalisches Ereignis *p*, beispielsweise der stechende Schmerz, den ich gerade in diesem Moment in meiner rechten Hand verspüre, verursacht eine Körper-

¹⁸⁰ Nicht-reduktive, physikalistische Ansätze sind durch folgende zwei Kernthesen gekennzeichnet: die Irreduzibilitätsthese und die Physikalismusthese. Erstere besagt, dass das Mentale nicht vollständig auf das Physikalische reduzierbar ist; wobei hier unter ‚Mentalem‘ im Wesentlichen mentale Eigenschaften verstanden werden (analog steht auch das Physikalische für physikalische Eigenschaften). Der Physikalismusthese zufolge gibt es nur eine Substanz, aus der alles Seiende besteht – und zwar die physikalische Substanz; jedes Einzelding ist also im Grunde physikalisch. Letzten Endes muss folglich zumindest jedes mentale Einzelereignis mit einem physikalischen Einzelereignis identisch sein. Im Hinblick auf Eigenschaften muss das aber nicht zwangsläufig so sein, es ist durchaus vorstellbar, dass mentale Eigenschaften nicht mit physikalischen identisch und folglich auch nicht auf solche reduzierbar sind. Ein anderer Sammelbegriff, der die wesentlichen Charakteristika von nicht-reduktiven, physikalistischen Theorien zusammenfasst, ist ‚Tokenphysikalismus‘.

¹⁸¹ Das Exklusionsproblem wird erstmals ansatzweise in einem Aufsatz aus dem Jahre 1968 von Norman Malcolm thematisiert – vgl. Malcolm, N.: „*The Conceivability of Mechanism*“ in: *Philosophical Review*, Vol. 77, 1968, S. 45-72. Kim selbst spricht meistens nicht vom Exklusionsproblem bzw. -argument, sondern vom ‚Supervenienz-Argument‘. Da aber in der Sekundärliteratur in der Regel vom ‚Exklusionsproblem‘ die Rede ist, verwende auch ich im Folgenden diesen Terminus.

¹⁸² Kim, Jaegwon: *Mind in a Physical World – An Essay on the Mind-Body Problem and Mental Causation*, Cambridge (Mass.), MIT Press, 2000, S. 37f.

bewegung, nämlich dass ich meinen rechten Arm zurückziehe. Gemeinhin nimmt man darüber hinaus an, dass die Kausalbeziehung zwischen m und p deshalb besteht, weil m der mentalen Art M (im Beispiel Schmerzempfindung) und p der physikalischen Art P (im Beispiel Körperbewegung) angehört. Nun stellt sich aber die Frage, ob p neben der mentalen Ursache m zugleich auch eine physikalische Ursache p^* besitzt, die der physikalischen Art N angehört. Im obigen Schmerzbeispiel könnte p^* etwa die Reizung gewisser C-Fasern in einem bestimmten Areal meines Gehirns sein und somit der Klasse der N -Ereignisse (neuronalen Ereignisse) angehören.

Zusammen mit zwei weiteren Prämissen illustriert dieses Szenario die Virulenz des Exklusionsproblems sehr deutlich. Diese beiden Vorannahmen besagen, dass erstens der Bereich der Physik kausal geschlossen ist und dass es zweitens keine systematische kausale Überdeterminierung im Falle von Geist-Körper-Wechselwirkungen gibt. Zunächst stellt sich die Frage, was diese beiden Zusatzannahmen genau bedeuten?

Die Forderung nach der kausalen Geschlossenheit des physikalischen Bereichs besagt in einer Ausformulierung, dass jedes physikalische Ereignis, das eine hinreichende Ursache hat, eine hinreichende *physikalische* Ursache besitzt. In anderen Worten kann es bei Kausalketten innerhalb der Physik keine Ursachen geben, die selbst nicht physikalisch sind, die also außerhalb der Physik zu suchen wären. Geht man in der kausalen Geschichte eines beliebigen physikalischen Wirkungsereignisses zurück (also jedes physikalischen Ereignisses, das *überhaupt* irgendeine Ursache hat), finden sich dort also *ausschließlich* andere physikalische Ereignisse und dementsprechend keine zu einer anderen Art zugehörigen Ereignisse. Die Welt der Physik ist daher sozusagen explanatorisch autark, d.h. für eine vollständige Erklärung von Kausalzusammenhängen zwischen physikalischen Ereignissen genügt ein Rekurs, der exklusiv auf physikalische Ereignisse Bezug nimmt. In diesem Sinne ist die kausale Geschlossenheit der physikalischen Welt zu verstehen.

Das Verbot der systematischen Überdeterminierung hebt auf den Umstand ab, dass eine Klasse von Ereignissen nicht zwei Arten von jeweils für sich alleine schon hinreichenden, voneinander unabhängigen Ursachen haben kann.¹⁸³ D.h. selbst wenn man in obigem Beispiel dazu tendiert anzunehmen, dass das Ereignis p (das Zurückziehen meines rechten Armes) *sowohl*

¹⁸³ Freilich könnte man geneigt sein einige *Sonderfälle* kausaler Überdeterminierung zuzulassen – hierzu ein kurzes Beispiel: In einem Haus entsteht im Wohnzimmer durch einen Kurzschluss ein Kabelbrand. Auf der anderen Seite des Hauses, im Schlafzimmer, schlägt gleichzeitig ein Blitz und entfacht ein Feuer. Das Haus brennt nun bis auf die Grundmauern nieder. In diesem Beispiel sind beide Ereignisse (Kabelbrand und Blitzschlag) jeweils für sich allein genommen schon hinreichend für das Abbrennen des ganzen Hauses – i.e. der Kabelbrand per se hätte zum Niederbrennen des Hauses geführt, auch wenn der Blitz nicht eingeschlagen hätte (und umgekehrt). Zudem sind Kabelbrand und Blitzschlag zwei voneinander unabhängige Ereignisse. In diesem Fall scheint ein Ereignis „Abbrennen des Hauses“ zwei voneinander unabhängige und jeweils für sich alleine schon hinreichende Ursachen zu haben und ergo kausal überdeterminiert zu sein.

eine hinreichende mentale Ursache m (den stechenden Schmerz) *als auch* eine hinreichende physikalische Ursache p^* (das C-Faserfeuern in meinem Gehirn) hat, wobei m und p^* jeweils für sich genommen schon für das Eintreten von p kausal *allein*verantwortlich sein könnten und vollkommen unabhängig voneinander sind, so erscheint es dennoch als sehr unwahrscheinlich, dass *alle* Ereignisse der Art P (Körperbewegung) eine hinreichende Ursache der Art M (Schmerzempfindung) *und gleichzeitig* eine hinreichende Ursache der Art N (Neuronenfeuern) haben. Denn es ist nicht ersichtlich warum eine Teilmenge von physikalischen Ereignissen (und zwar diejenigen physikalischen Ereignisse, die auch eine mentale Ursache haben) systematisch überdeterminiert sein sollte, während eine andere Teilmenge physikalischer Ereignisse nur einfach (nämlich durch andere physikalische Ereignisse) determiniert ist. Es sei denn man wäre bereit einzugestehen, dass es sich in beiden Fällen um unterschiedliche Arten von Kausalität handelt – einerseits rein physikalische Kausalität und andererseits um psychophysische Kausalität – die ihren jeweils eigenen Gesetzen folgen. Dies wiederum wäre aber offensichtlich ein ‚Rückfall‘ in den interaktionistischen Substanzdualismus und brächte die entsprechenden Probleme mit sich. Zudem müssten weiterhin die nicht zu verleugnenden Korrelationen zwischen physikalischen Ereignissen und mentalen Ereignissen erklärt werden. Gehen etwa Schmerzempfindungen immer mit einem bestimmten Muster von Neuronenfeuern im Gehirn des Schmerzempfindenden einher und folgt dem Schmerz immer ein bestimmtes anderes physikalisches Ereignis (etwa eine Körperbewegung), dann scheint es als wäre hier sowohl psychophysische Kausalität (die Schmerzempfindung verursacht eine bestimmte Körperbewegung), als auch rein physikalische Kausalität (das neuronale Ereignis verursacht eine Körperbewegung) am Werke – und dies führt wiederum zur Annahme einer systematischen Überdeterminierung, die man gerade durch das Einführen von zwei Grundarten von Kausalität vermeiden wollte.

Ist nun aber die obige Beschreibung der kausalen Geschichte des Schmerzes in meiner rechten Hand kohärent und sind die beiden zusätzlichen Prämissen wahr, dann steht man vor folgendem Dilemma¹⁸⁴: Entweder ist man gezwungen zu argumentieren, dass das physikalische Ereignis p (meine Armbewegung) neben der mentalen Ursache m (stechender Schmerz) *keine weitere* (und insbesondere keine physikalische) Ursache hat oder aber man müsste zugeben, dass es *neben* m eine *weitere* hinreichende Ursache p^* (C-Faserfeuern) für das Eintreten von p gibt. Beide Gedankengänge müsste man jedoch zurückweisen, wenn man an den der kausalen Geschlossenheit der Physik und an dem Verbot einer systematischen Überdeterminierung

¹⁸⁴ Ein Ausweg aus dem Dilemma bestünde natürlich darin zu erklären m ist identisch mit p^* oder m ist vollständig reduzierbar auf p^* . Es ist jedoch offensichtlich, dass einem nicht-reduktiven Physikalisten eine solche Antwortstrategie aufgrund seines Credos der Nicht-Reduzierbarkeit des Mentalen verwehrt bleiben muss.

festhalten will. Schließlich ist die erste Argumentationsweise ein offensichtlicher Verstoß gegen die kausale Geschlossenheit der Physik und der zweite Gedankengang würde eine (systematische) Überdeterminierung zulassen. Egal welchen Weg man also einschlägt, es scheint als könnte man nicht gleichzeitig auf der obigen Beschreibung der Kausalbeziehung zwischen meinem Schmerz und meiner Armbewegung und den Ideen der kausalen Geschlossenheit der Physik und dem Verbot systematischer Überdeterminierung beharren. Entweder lässt man die Vorstellung einer kausal geschlossenen Physik fallen und nimmt an, dass das Auftreten eines mentalen Ereignisses *allein* schon kausal hinreichend für das Eintreten eines physikalischen Ereignisses sein könnte oder aber man lässt *zwei* unabhängige, per se jeweils schon hinreichende Ursachen für ein physikalisch Ereignis zu – ein mentales und ein anderes physikalisches Ereignis. Beide ‚Lösungen‘ sind jedoch äußerst unbefriedigend, weil sie gegen je eine intuitiv sehr einleuchtende Hypothese verstoßen.

An späterer Stelle präsentiert Kim in *Mind in a Physical World* das Exklusionsproblem ausführlich als Argument in Form eines Dilemmas:

I will now proceed to construct a dilemma-style argument that apparently leads to the conclusion that mental causation is unintelligible. In essence the argument to be presented is the result of superimposing mind-body supervenience on the causal exclusion problem. We begin by setting forth the two horns of the dilemma:

- (i) Either mind-body supervenience holds or it fails.
 - (...)
 - Mind-body supervenience* Mental properties supervene on physical properties in the sense that if something instantiates any mental property *M* at *t*, there is a physical base property *P* such that the thing has *P* at *t*, and necessarily anything with *P* at a time has *M* at that time.
 - (...)
 - (ii) If mind-body supervenience fails, there is no visible way of understanding the possibility of mental causation. (...)
 - (iii) Suppose that an instance of mental property *M* causes another mental property *M** to be instantiated. (...)
 - (iv) *M** has a physical supervenience base *P**. (...)
 - (v) *M** is instantiated on this occasion: (a) because, ex hypothesi, *M* caused *M** to be instantiated; (b) because *P**, the physical supervenience base of *M**, is instantiated on this occasion. (...)
 - (vi) *M* caused *M** by causing *P**. That is how this instance of *M* caused *M** to be instantiated on this occasion. (...)
 - (vii) *M* itself has a physical supervenience base *P*. (...)
 - (viii) *P* caused *P**, and *M* supervenes on *P* and *M** supervenes on *P**. (...)
 - (ix) The *M*-to-*M** and *M*-to-*P** causal relations are only apparent, arising out of a genuine causal process from *P* to *P**.
- Whence a dilemma:
- (x) If mind-body supervenience fails, mental causation is unintelligible; if it holds, mental causation is again unintelligible. Hence mental causation is unintelligible.¹⁸⁵

¹⁸⁵ Kim, Jaegwon: *Mind in a Physical World – An Essay on the Mind-Body Problem and Mental Causation*, Cambridge (Mass.), MIT Press, 2000, S. 39-46.

An den Ausgangspunkt für das Exklusionsargument stellt Kim die beiden Hörner des Dilemmas – das erste Horn ist ‚Geist-Körper-Supervenienz trifft zu‘ und dementsprechend lautet das zweite Horn ‚Geist-Körper-Supervenienz trifft nicht zu‘.

1. Entweder trifft Geist-Körper-Supervenienz zu oder sie trifft nicht zu.

Kim definiert ‚Geist-Körper-Supervenienz‘ in diesem Zusammenhang folgendermaßen: Mentale Eigenschaften supervenieren in dem Sinne auf physikalischen Eigenschaften, dass, wenn irgendetwas eine mentale Eigenschaft M zu einem Zeitpunkt t instanziiert, es eine physikalische Basiseigenschaft P gibt, so dass gilt: jedes Ding, das P zum irgendeinem Zeitpunkt hat, hat notwendigerweise zu diesem Zeitpunkt auch M. Bei der Notwendigkeit, von der in diesem Kontext die Rede ist, handelt es sich (mindestens) um nomologische Notwendigkeit. D.h. wenn Geist-Körper-Supervenienz in diesem Falle zutrifft, trifft sie auch in allen möglichen Welten zu, in denen dieselben Naturgesetze gelten wie in der aktuellen Welt.

In der zweiten Prämisse seines Arguments analysiert Kim kurz das zweite Horn des Dilemmas. Er stellt sich die Frage was wäre, wenn Geist-Körper-Supervenienz nicht zuträfe und stellt bei deren Beantwortung folgende These auf:

2. Wenn Geist-Körper-Supervenienz nicht zutrifft ist mentale Verursachung unverstän-
lich.

Die Intelligibilität von mentaler Verursachung hängt für Kim also wesentlich mit dem Vorhandensein von Geist-Körper-Supervenienz zusammen und in der Regel teilen die meisten Physikalisten diese Ansicht. Der Grund hierfür besteht darin, dass die kausale Geschlossenheit der Physik für sie als unabdingbar gilt. Die Kernaussage des Physikalismus besagt nun gerade, dass es nur eine Substanz gibt und zwar die physikalische. Insofern überrascht es nicht weiter, dass Physikalisten die kausale Geschlossenheit der Physik als gegeben voraussetzen. Wenn der Bereich der Physik hingegen kausal offen wäre, müssten entweder nicht-physikalische Ursachen zugelassen werden oder es müsste physikalische Ereignisse geben, die nicht verursacht wurden. Ersteres wäre sicherlich mit keiner Spielart des Physikalismus vereinbar, denn diese Annahme käme offensichtlich einem ontologischen Postulat gleich, das der zentralen Aussage des Physikalismus diametral gegenüber steht. Zweiteres wiederum ist kaum mit einer physikalistischen Auffassung von Kausalität vereinbar. Die Forderung nach der unabhängigen Initiierung von Kausalketten erinnert doch sehr stark an eine substanzduali-

stische Auffassung von Verursachung. Die kausale Geschlossenheit der Physik definiert Kim so: „One way of stating the principle of physical causal closure is this: If you pick any physical event and trace out its causal ancestry or posterity, that will never take you outside the physical domain. That is, no causal chain will ever cross the boundary between the physical and the nonphysical.“¹⁸⁶ Wer die kausale Geschlossenheit der Physik anzweifelt, muss der Physik gleichzeitig auch die prinzipielle Möglichkeit zur Vollständigkeit absprechen. In diesem Fall wäre es nämlich nicht mehr ausnahmslos möglich jedes physikalische Phänomen allein durch Rekurs auf andere physikalische Ereignisse zu erklären. Vor diesem Hintergrund ist es einfach zu verstehen, warum mentale Verursachung ohne Geist-Körper-Supervenienz unverständlich wird: „Now if mind-body supervenience fails – that is, if the mental domain floats freely, unanchored in the physical domain, causation from the mental to the physical would obviously breach the physical causal closure.“¹⁸⁷ Schließlich ist der Geist-Körper-Supervenienz zufolge jedes mentale Phänomen im Bereich der Physik verankert. Daher muss es auch eine Menge von physikalischen Bedingungen geben, die für das Auftreten eines bestimmten mentalen Phänomens nomologisch hinreichend sind und von denen somit dessen Eintreten abhängt. Jedes mentale Ereignis bedarf deswegen einer physikalischen Basis. Nur auf diese Weise kann das Mentale in den Wirkungskreis des Physikalischen rücken. Folglich ist auch klar, dass das Mentale nicht als ontologisch autonomer Bereich existieren kann und deshalb auch keine eigenständigen kausalen Einflüsse ausüben kann.

Während einige Physikalisten Geist-Körper-Supervenienz als Patentlösung für das Verständnis mentaler Verursachung im Physikalismus preisen, ist Kim demgegenüber eher skeptisch eingestellt: „(...) mind-body supervenience, far from being part of the solution, (...) may turn out to be part of the problem.“¹⁸⁸ Kim glaubt also, dass Geist-Körper-Supervenienz nicht, wie von vielen erhofft, der große Heilsbringer für den Physikalismus ist, sondern durchaus auch mit Problemen behaftet ist. Um diesen Gedanken weiterzuspinnen, macht Kim in der dritten Prämisse folgende Annahme:

3. Angenommen die Instanziierung einer mentalen Eigenschaft M verursacht die Instanziierung einer anderen mentalen Eigenschaft M*.

¹⁸⁶ Kim, Jaegwon: *Mind in a Physical World – An Essay on the Mind-Body Problem and Mental Causation*, Cambridge (Mass.), MIT Press, 2000, S.40.

¹⁸⁷ Kim, Jaegwon: *Mind in a Physical World – An Essay on the Mind-Body Problem and Mental Causation*, Cambridge (Mass.), MIT Press, 2000, S.40.

¹⁸⁸ Kim, Jaegwon: *Mind in a Physical World – An Essay on the Mind-Body Problem and Mental Causation*, Cambridge (Mass.), MIT Press, 2000, S.41.

Kims Annahme an dieser Stelle ist in der Tat ganz unschuldig – er geht lediglich von einem Fall von Geist-zu-Geist-Verursachung aus, wie man ihn tagtäglich dutzende Male bei sich selbst erleben kann: die Instanziierung einer mentalen Eigenschaft, etwa eines Durstgefühls, verursacht die Instanziierung einer anderen mentalen Eigenschaft, beispielsweise den Wunsch nach einem Getränk. Zugleich muss aber für die (verursachende) mentale Eigenschaft aufgrund der These der Geist-Körper-Supervenienz gelten:

4. M^* hat die physikalische Supervenienzbasis P^* .

Der These der Supervenienz zufolge muss nämlich jede Instanziierung einer mentalen Eigenschaft mit dem gleichzeitigen Auftreten einer physikalischen Supervenienzbasis einhergehen. Im Beispiel etwa würde ein Durstgefühl immer dann (und auch nur dann) entstehen, wenn bestimmte Rezeptoren im Hypothalamus in gewisser Weise auf sie umgebende Umstände reagieren. Allerdings steht man dann unmittelbar vor folgendem Problem – die mentale Wirkungseigenschaft scheint zwei Ursachen zu besitzen: das Durstgefühl einerseits und die Reaktion gewisser Rezeptoren im Hypothalamus auf deren unmittelbares Milieu andererseits. So gelangt Kim zur fünften Prämisse:

5. M wird zu diesem Zeitpunkt instanziiert, weil (a) ex hypothesi gilt: die Instanziierung von M verursacht die Instanziierung von M^* , oder weil (b) P^* , die physikalische Supervenienzbasis von M^* , zu diesem Zeitpunkt instanziiert wird.

Plötzlich scheint es als gäbe es zwei mögliche Ursachen für die Instanziierung von M^* (dem Wunsch nach einem Getränk): einerseits könnte man annehmen M (das Durstgefühl) habe M^* verursacht. Andererseits könnte man aber auch geneigt sein den Grund für die Instanziierung von M^* in P^* (der Reaktion bestimmter Rezeptoren im Hypothalamus auf ihre Umwelt) zu sehen. Zwischen diesen beiden Kausalerklärungen für M^* sieht Kim ein starkes Spannungsverhältnis. Aufgrund der Supervenienzthese tritt M^* nämlich immer genau dann ein, wenn P^* eintritt – und zwar unabhängig davon, was vorher passiert ist. Somit scheint es also auch keine Rolle zu spielen, ob M vor M^* eingetreten ist oder nicht. Ergo wird der kausale Status von M als Ursache von M^* durch P^* akut bedroht. Allein die Instanziierung von P^* ist schon hinreichend für das Eintreten von M^* oder in anderen Worten um M^* zu verursachen. Die einzige verbleibende Möglichkeit wie etwas überhaupt mit der Verursachung von M^* in Verbin-

dung gebracht werden kann, besteht also über die Beziehung zu M*s Supervenienzbasis P*. Über diesen Gedankengang gelangt Kim zur sechsten Prämisse:

6. M verursacht M* indem es P* verursacht. Auf diese Weise hat die Instanziierung von M zu diesem Zeitpunkt die Instanziierung von M* verursacht.

Die einzige Möglichkeit wie M für das Auftreten von M* kausal verantwortlich sein kann, so Kim, besteht darin, dass M P* verursacht. Dieser Vorstellung liegt laut Kim folgendes Prinzip zugrunde: „*To cause a supervenient property to be instantiated, you must cause its base property (or one of its base properties) to be instantiated.*“¹⁸⁹ Eine kurze Illustration an einem Beispiel verdeutlicht dies: Um sich etwa seiner Kopfschmerzen zu entledigen, muss man beispielsweise eine Aspirin (oder eine andere Tablette mit dem Wirkstoff Acetylsalicylsäure) einnehmen.¹⁹⁰ D.h. man muss kausal in die Hirnprozesse eingreifen, auf denen die Kopfschmerzen supervenieren. Dies ist auch der *einzige* Weg, der beschritten werden kann, um die Kopfschmerzen zu beseitigen.

Die Forderung, dass M M* verursacht, indem es P* verursacht, beinhaltet interessanterweise einen Ebenenwechsel. Während man anfangs von einem Fall von Geist-zu-Geist-Verursachung ausgegangen ist, wird an dieser Stelle die Grenze zum physikalischen Bereich überschritten. Es ist offensichtlich, dass es sich um eine Geist-zu-Körper-Verursachung handelt, wenn M P* verursacht. Daher gelangt Kim zu dem aufschlussreichen Fazit: „*Under the mind-body supervenience assumption, mental-to-mental causation implies, or presupposes, mental-to-physical causation.*“¹⁹¹

Der nächste Schritt von Kims Argument besteht in einer simplen Anwendung der Supervenienzthese auf M:

7. M selbst hat eine physikalische Supervenienzbasis – und zwar P.

An dieser Stelle droht ein erneuter Wettbewerb um die Verursachung von P* auszubrechen – und zwar zwischen M und P. Allerdings ist es naheliegend, dass man zu einem ähnlichen Schluss wie im ersten Fall gelangt. Vergleicht man nun nämlich M und P hinsichtlich ihres

¹⁸⁹ Kim, Jaegwon: *Mind in a Physical World – An Essay on the Mind-Body Problem and Mental Causation*, Cambridge (Mass.), MIT Press, 2000, S.42.

¹⁹⁰ Selbstverständlich gibt es auch noch andere Wege kausal in die Kopfschmerzen verursachenden Hirnprozesse einzugreifen – die Einnahme einer Schmerztablette soll an dieser Stelle nur exemplarisch als eine von vielen Möglichkeiten verstanden werden.

¹⁹¹ Kim, Jaegwon: *Mind in a Physical World – An Essay on the Mind-Body Problem and Mental Causation*, Cambridge (Mass.), MIT Press, 2000, S.43.

kausalen Status im Hinblick auf P^* , so scheint es als würde P M seiner kausalen Kraft berauben. Denn sieht man Kausalität in nomologischer Notwendigkeit begründet, dann wäre es durchaus legitim P als Ursache von P^* zu betrachten. Schließlich ist P für sich alleine schon hinreichend für P^* , weil P für M hinreichend ist und M wiederum hinreichend für P^* ist. Betrachtet man Kausalität hingegen in Sinne kontrafaktischer Abhängigkeit, so qualifiziert sich P auch unter diesen Umständen als Ursache von P^* . Wenn nämlich P nicht eingetreten wäre, wäre damit auch M nicht eingetreten (gesetzt des Falles, dass es zu diesem Zeitpunkt keine andere physikalische Basis für M gegeben hatte). Falls aber M nicht eintritt, tritt auch P^* nicht ein. Somit kann man mit Fug und Recht behaupten, dass wenn P nicht eingetreten wäre, auch P^* nicht eingetreten wäre. Nun aber scheint ein Überschuss an Ursachen vorzuherrschen, denn sowohl M als auch P scheinen für sich alleine jeweils schon hinreichende Ursachen von P^* zu sein. Die Bedrohung einer kausalen Überdeterminierung kann *nicht* dadurch abgewandt werden, dass eine Kausalkette P -zu- M -zu- P^* postuliert wird. Es wäre nämlich äußerst unglücklich ein Supervenienzverhältnis als Kausalbeziehung zu begreifen, weil erstens die Instanziierungen der Eigenschaften in einer Supervenienzbeziehung gleichzeitig stattfindet wohingegen in einer Kausalbeziehung die Ursache immer der Wirkung zeitlich vorangeht. Zweitens ist es völlig unklar was kausal zwischen die subvenienten und supervenienten Eigenschaften geschaltet sein sollte und ob dies dann selbst mental oder physikalisch ist. Zudem würde, wenn M die Ursache von P^* wäre, wiederum das Prinzip der Geschlossenheit der Physik verletzt. Genauso wenig könnte man M und P zusammen als einzelne Ursache von P^* ansehen. Schließlich sind P und M schon jeweils alleine hinreichend für P^* , daher ist nicht ersichtlich warum P und M gemeinsam mehr kausale Kraft haben sollten als nur P oder M für sich genommen. Zudem würde dieses Mehr an Kausalkraft auch gar nicht benötigt um P^* zu verursachen – schließlich sind P und M jeweils für sich alleine kausal hinreichend für das Eintreten von M^* . Ferner ist der Ansatz, der in M und P gemeinsam eine Ursache von P^* sieht, abermals nicht mit dem Geschlossenheitsprinzip kompatibel. Dieser Ansatz wäre nämlich nur dann plausibel, wenn M als eigenständiger und notwendiger Bestandteil in der Verursachung von P^* betrachtet werden würde. Damit würde aber erneut ein nicht-physikalisches Element in die Kausalerklärung von P^* treten. Schlussendlich kann auch nicht argumentiert werden, dass es sich um einen Fall kausaler Überdeterminierung handelt. Abgesehen davon, dass dies die sehr unplausible Folge hätte, dass somit *jeder* Fall von mentaler Verursachung kausal überdeterminiert wäre, gäbe es zwei weitere Schwierigkeiten: zum Ersten könnte man den Eindruck gewinnen als würden mentale Ursachen überflüssig gemacht, indem ihnen eine hinreichende physikalische Ursache an die Seite gestellt wird. Zweitens wäre dann abermals

die Geschlossenheit der Physik in Gefahr. Hierzu stelle man sich eine mögliche Welt vor, die in möglichst vielen Belangen der aktuellen Welt gleicht, in der aber die mentale Ursache eintritt, während die physikalische Ursache ausbleibt. In diesem Fall besäße das physikalische Wirkungsereignis nur noch eine einzige Ursache – und zwar die mentale; womit wiederum aber das Geschlossenheitsprinzip in Frage gestellt werden würde. Eine für Kim unzulässige Konsequenz: „I do not think we can accept this consequence: that a minimal counterfactual supposition like that can lead to a major change in the world.“¹⁹² Für Kim kumuliert die natürlichste Form der Rekonstruktion dieser Situation in folgender Aussage:

8. P hat P* verursacht und M superveniert auf P und M* superveniert auf P*.

Durch dieses Szenario werden einerseits die Regelmäßigkeiten zwischen Instanziierungen von M und Instanziierungen von M*, sowie zwischen Instanziierungen von M und Instanziierungen von P*, erklärt. Diese Regelmäßigkeiten kommen freilich nicht rein zufällig zu Stande. Allerdings unterscheidet dieses Bild der Situation eindeutig zwischen den *wirklichen* Kausalbeziehungen und zwischen Regelmäßigkeiten, die sozusagen ‚parasitär‘ auf den wahren Kausalprozessen haften. Für Kim ist die Ursache-Wirkungsbeziehung zwischen den mentalen Eigenschaften M und M* ausschließlich *epiphenomenal* zu verstehen. In diesem Zusammenhang stellt er folgende Analogie vor: „In the case of the supposed *M-M** causation, the situation is rather like a series of shadows cast by a moving car: there is no causal connection between the shadow of the car at one instant and its shadow an instant later, each being an effect of the moving car. The moving car represents a genuine causal process, but the series of shadows it casts, however regular and lawlike it may be, does not constitute a causal process.“¹⁹³ Alles in allem gelangt Kim daher zu folgender Schlussfolgerung:

9. Die Kausalbeziehungen M-zu-M* und M-zu-P* existieren nur scheinbar, sie entstehen nur auf Basis des einzig wirklichen Kausalprozesses P-zu-P*.

Unmittelbar daraus leitet Kim auch ein Dilemma ab:

¹⁹² Kim, Jaegwon: *Mind in a Physical World – An Essay on the Mind-Body Problem and Mental Causation*, Cambridge (Mass.), MIT Press, 2000, S.45.

¹⁹³ Kim, Jaegwon: *Mind in a Physical World – An Essay on the Mind-Body Problem and Mental Causation*, Cambridge (Mass.), MIT Press, 2000, S.45.

10. Wenn Geist-Körper-Supervenienz nicht zutrifft, ist mentale Verursachung nicht verständlich. Wenn Geist-Körper-Supervenienz zutrifft, ist mentale Verursachung auch nicht verständlich. Also ist mentale Verursachung nicht verständlich.

Zum Abschluss des Arguments bilanziert Kim:

That is then the supervenience argument against mental causation, or Descartes's revenge against the physicalists. I believe that it poses a serious challenge to physicalism by casting doubts on the possibility of mental causation within the parameters it sets for itself. Descartes's difficulties arose from the duality of mental and material substances. Current mainstream physicalism, which calls itself "nonreductive physicalism," runs into parallel difficulties on account of its commitment to the duality of psychological and physical properties – or its failure to make a reductionist commitment for psychological properties. For it is clear that the tacit assumption that gets the supervenience argument going is mind-body antireductionism; if mental properties are viewed as reducible to physical properties in an appropriate way, we should expect to be able to disarm the argument (although of course the details will need to be worked out).¹⁹⁴

Das Exklusionsargument illustriert laut Kim also Parallelen zwischen den Problemen, die Descartes' interaktionistischer Substanzdualismus und der sogenannte nicht-reduktive Physikalismus mit einer adäquaten Erklärung mentaler Verursachung haben. Während die Probleme für eine kohärente Darstellung mentaler Verursachung bei Descartes in der Dualität der mentalen und der physikalischen Substanz begründet sind, resultieren sie beim nicht-reduktiven Physikalismus aus der Dualität psychologischer und physikalischer Eigenschaften. Bestünde jedoch die Möglichkeit einer angemessenen Reduktion psychologischer Eigenschaften auf physikalische Eigenschaften, dann könnte man auch das Exklusionsproblem als aufgelöst betrachten.

In seinem Buch *Physicalism or something near enough* verfeinert Kim die Struktur des Exklusionsproblems noch weiter, indem er einige Prinzipien herausarbeitet, die vermutlich die meisten (nicht-reduktiven Physikalisten) als wahr anerkennen würden, die aber zusammengekommen mentale Verursachung im nicht-reduktiven Physikalismus unmöglich zu machen scheinen. Kim benennt und definiert die Prinzipien folgendermaßen:

The causal closure of the physical domain. If a physical event has a cause at t , then it has a physical cause at t .
(...)

Principle of causal exclusion. If an event e has a sufficient cause c at t , no event at t distinct from c can be a cause of e (unless this is a genuine case of causal overdetermination).
(...)

Principle of determinative/generative exclusion. If the occurrence of an event e , or an instantiation of a property P , is determined/generated by an event c – causally or otherwise – then e 's occurrence is not determined/generated by an event wholly distinct from or independent of c – unless this is a genuine case of overdetermination.

¹⁹⁴ Kim, Jaegwon: *Mind in a Physical World – An Essay on the Mind-Body Problem and Mental Causation*, Cambridge (Mass.), MIT Press, 2000, S. 46.

(...)

The problem of mental causation. Causal efficacy of mental properties is inconsistent with the joint acceptance of the following four claims: (i) physical causal closure, (ii) causal exclusion, (iii) mind-body supervenience, and (iv) mental/physical property dualism – the view that mental properties are irreducible to physical properties.¹⁹⁵

An dieser Stelle verdeutlicht Kim nochmals die Grundannahmen, auf denen das Exklusionsproblem beruht und erklärt wie deren Zusammenwirken zur Entstehung des Problems führt.

Das Prinzip der kausalen Geschlossenheit der Physik bekundet, dass, wenn ein physikalisches Ereignis zu einem Zeitpunkt *t* eine Ursache hat, es zum Zeitpunkt *t* eine *physikalische* Ursache hat.¹⁹⁶ Worauf Kim bei dieser Definition besonderes Augenmerk legt ist, dass der Bereich der Physik dadurch sowohl kausal als auch explanatorisch *selbstgenügsam* bzw. *autark* ist. Das bedeutet, dass man für das Auffinden einer adäquaten und vollständigen Kausalerklärung für ein physikalisches Wirkungsereignis den Bereich der Physik nicht verlassen muss. Offensichtlich ist dieses Verständnis der kausalen Geschlossenheit der Physik nicht äquivalent mit der Auffassung, dass es ausschließlich physikalische Entitäten gibt oder dass die physikalische Kausalität die einzig existente oder legitime Form von Verursachung darstellt. Im Gegenteil, diese Vorstellung von der kausalen Geschlossenheit der Physik ist im Grunde durchaus auch mit (substanz)dualistischem Gedankengut vereinbar – allerdings nur solange keine kausale Wechselwirkung zwischen physikalischen und nicht-physikalischen Ereignissen oder Entitäten postuliert wird. Dementsprechend fällt Descartes' interaktionistischer Substanzdualismus dieser Auflage zum Opfer, wohingegen sich etwa Leibniz' prästabilierte Harmonie als damit verträglich erweisen würde. Ferner gilt es zu beachten, dass die Idee der Vollständigkeit der Physik nicht bereits nicht-physikalische Ursachen aus Kausalerklärungen für physikalische Ereignisse ausschließt.¹⁹⁷ Dies geschieht erst durch das zweite Prinzip – das Prinzip der kausalen Exklusion. Diesem Herzstück des Exklusionsarguments zufolge kann ein Ereignis *e*, das zum Zeitpunkt *t* eine hinreichende Ursache *c* hat, kein anderes Ereignis als das Ereignis *c* zum Zeitpunkt *t* zur Ursache haben, es sei denn es handelt sich um einen Fall genuiner kausaler Überdeterminierung. Laut Kim ist diese Ausformulierung des Prinzips der kausalen Exklusion ein generelles, metaphysisches Prinzip, das sich nicht unbedingt auf mentale und physikalische Ursachen beziehen muss und das sich zwischen Mentalem und Physikalischem sozusagen neutral verhält (d.h. keine der beiden Kausalerklärungen als besser als die

¹⁹⁵ Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 15-22.

¹⁹⁶ Ein Analogon zu dieser Definition der kausalen Geschlossenheit der Physik lautet nach Kim wie folgt: Wenn ein physikalisches Ereignis eine kausale Erklärung hat, hat es eine physikalische kausale Erklärung.

¹⁹⁷ Vgl. Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 15ff.

andere einstuft).¹⁹⁸ Für das Exklusionsargument formuliert Kim eine generalisierte Version des Prinzips der kausalen Exklusion – und zwar das Prinzip der determinativen / generativen Exklusion. Dessen Definition lautet: Wenn das Vorkommen eines Ereignisses *e* oder die Instanziierung einer Eigenschaft *P* durch ein Ereignis *c* kausal oder anders determiniert ist oder erzeugt wird, dann ist *e*'s Vorkommen nicht durch ein Ereignis, das gänzlich verschieden bzw. unabhängig von *c* ist, determiniert bzw. erzeugt – es sei denn es handelt sich um einen Fall genuiner kausaler Überdeterminierung. Diese generalisierte Version des Prinzips der kausalen Exklusion weitet Kausalität oder kausale Determinierung zu Determinierung bzw. Erzeugung simpliciter aus, also unabhängig davon ob diese Determinierung oder Erzeugung kausal ist oder nicht. Die intuitive Idee hinter diesem Prinzip ist, dass ein Ereignis oder die Instanziierung einer Eigenschaft ihre Existenz allein einem anderen Ereignis verdankt, es sozusagen daraus hervorgebracht wird oder in anderen Worten sich seine Existenz direkt daraus ableiten lässt. In diesem Sinne wird Kausalität also als Determination oder Erzeugung aufgefasst – diese Vorstellung von Verursachung ist, laut Kim, in vielen Aspekten stärker (in mancherlei Hinsicht aber auch schwächer) als die konkurrierende Auffassung von Kausalität als kontrafaktischer Abhängigkeit. Zudem sei es Kausalität in eben jenem Sinne, die bei Problemen mentaler Verursachung gemeinhin eine gewichtige Rolle spiele. Nach Kim ist Supervenienz eine andere Art einen Zustand oder die Instanziierung einer Eigenschaft zu erzeugen: so werden nach diesem Verständnis etwa die ästhetischen Eigenschaften eines Kunstwerks von seinen physikalischen Eigenschaften generiert.¹⁹⁹ D.h. um etwa die ästhetischen Eigenschaften eines Bildes zu verändern – i.e. um es z.B. schöner zu machen – müssten die physikalischen Eigenschaften des Bildes verändert werden: ohne eine Veränderung der physikalischen Eigenschaften ist eine Verschönerung des Bildes unmöglich.

Nach Kim ist es einfach zu sehen, wie die oben genannten Prinzipien für Verfechter einer Geist-Körper-Supervenienz, also für jeden, der eine Minimalform des Physikalismus vertritt, große Schwierigkeiten hinsichtlich der Vorstellbarkeit mentaler Verursachung bereiten. Diesen Bedenken hat Kim in seinem sogenannten ‚Supervenienzargument‘²⁰⁰ Ausdruck verliehen:

¹⁹⁸ Vgl. Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 17.

¹⁹⁹ Vgl. Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 17f.

²⁰⁰ In der Literatur ist gewöhnlich vom Exklusionsargument die Rede. Dazu Kim: „For usage of uniformity, it is best to think of the supervenience argument as a special form of the exclusion argument, and take the latter as a generic form of argument with the conclusion that mental cause is always excluded by physical cause.“ (in: Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 19.)

Briefly, the argument goes like this. Suppose that an instantiation of mental property M causes another mental property, M*, to instantiate. (We take property instantiations as events; instantiations of a mental property are mental events, and similarly for physical properties and physical events.) This is perfectly consistent with physical causal closure. But mind-body supervenience says that this instantiation of mental property M* occurs in virtue of the fact that one of the physical properties on which M* supervenes is instantiated at that time; call this physical base property P*. This means that given that P* is instantiated on this occasion, M* must of necessity be instantiated on this occasion. That is, the M*-instance is wholly dependent on, and is generated by, the P*-instance. At this point, the exclusion principle kicks in: Is the occurrence of the M*-instance due to its supposed cause, the M-instance, or its supervenience base event, P*-instance? It must be one or the other, but which one? Given the physical supervenience base P* is instantiated on this occasion, M* must be instantiated as well on this occasion, regardless of what might have preceded this M*-instance. In what sense, then, can the M-instance be said to be a “cause”, or a generative source, of the M*-instance?²⁰¹

Wenn die Instanziierung einer mentalen Eigenschaft M die Instanziierung einer anderen mentalen Eigenschaft M* verursacht, muss ein Vertreter von Geist-Körper-Supervenienz annehmen, dass die Instanziierung von M* deshalb eintritt, weil eine physikalische Eigenschaft P*, auf der M* superveniert, instanziiert wird. Bei jeder Gelegenheit, bei der P* instanziiert ist, muss also auch M* *notwendigerweise* instanziiert sein. Dem Exklusionsprinzip zufolge stellt sich nun aber die Frage, ob das Eintreten von M* der zunächst angenommenen Ursache M oder deren physikalischer Supervenienzbasis P* zuzurechnen ist. Da die Supervenienzthese der Basis P* den Primat in der Erklärung der Verursachungsrelation einräumt, stellt sich die Frage in welchem Sinn man überhaupt noch davon sprechen kann, dass die Instanziierung von M tatsächlich die Instanziierung von M* verursacht hat. Kim sieht die einzig plausible Lösung darin: „ (...) the M-instance caused the M*-instance *by causing* the P*-instance. More generally, the following principle seems highly plausible: *In order to cause a supervenient property to be instantiated, you must cause one of its base properties to be instantiated.*“²⁰² Die Instanziierung von M hat also die Instanziierung von M* verursacht, *indem* sie die Instanziierung von P* verursacht hat. Nach der Vorstellung der Geist-Körper-Supervenienz muss aber nun auch die vermeintliche Ursache M eine physikalische Supervenienzbasis besitzen, etwa P. Im Großen und Ganzen spricht dann jedoch alles dafür, P auch als die Ursache von P* zu betrachten: versteht man Kausalität als hinreichende Bedingung, so ist P eindeutig hinreichend für P* (weil P hinreichend für M ist, welches wiederum hinreichend für M* ist), versteht man Kausalität hingegen als kontrafaktische Abhängigkeit, so wäre P* nicht eingetreten, wenn P nicht eintritt (weil ohne P würde auch M nicht eintreten, was zur Folge hätte, dass auch M* nicht eintritt). Somit hat man folglich *zwei* kausale Behauptungen – und zwar ‚M verursacht P*‘ und ‚P verursacht P*‘. Der psychophysische Eigenschaftsdualismus, den ein nicht-reduktiver Physikalist per definitionem vertreten muss (eben um nicht-reduktiver Physikalist zu sein), zwingt zu der Sichtweise, dass M und P verschiedene Eigenschaften sind. Da-

²⁰¹ Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 19.

²⁰² Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 20.

her hat P* auch zwei unterschiedliche, jeweils für sich alleine schon hinreichende Ursachen, die zum selben Zeitpunkt eintreten. Nach dem Exklusionsprinzip muss aber entweder P oder M als Ursache von P* ausgeschlossen werden. Vernünftigerweise muss man nun M als Ursache ausschließen, denn wenn P ausgeklammert werden würde, käme das Prinzip der kausalen Abgeschlossenheit der Physik ins Spiel. Demgemäß gilt, wenn das physikalische Ereignis P* eine Ursache hat, muss es eine physikalische Ursache haben. Da M nun aber nicht physikalisch ist und auch nicht sein kann – schließlich kann M aufgrund der Annahme der Nicht-Reduzierbarkeit nicht mit P identisch sein – kommt es nicht als P*s Ursache in Frage.²⁰³

Vor dieser Kulisse entsteht das oben bereits zitierte Problem mentaler Verursachung. Die kausale Wirksamkeit mentaler Eigenschaften ist nicht vereinbar mit den folgenden vier Behauptungen: erstens der kausalen Geschlossenheit der Physik, zweitens dem Prinzip der kausalen Exklusion, drittens der Vorstellung von Geist-Körper-Supervenienz, und viertens der Irreduzibilität mentaler Eigenschaften auf physikalische Eigenschaften. Da alle Physikalisten die kausale Geschlossenheit der Physik und die Geist-Körper-Supervenienz verteidigen dürfen und da das Exklusionsprinzip eine allgemeingültige, metaphysische Einschränkung ist, kann die Schwachstelle nach Kim offensichtlich nur in der Irreduzibilität mentaler Eigenschaften liegen. Darin sieht er die einzige Stellschraube in diesem Argument, an der noch gedreht werden kann ohne die Grundgedanken des Physikalismus preiszugeben oder das augenscheinlich richtige Exklusionsprinzip leugnen zu müssen. Dementsprechend fordert Kim auch die These des Eigenschaftsdualismus fallen zu lassen und stattdessen einen reduktionistischen Standpunkt einzunehmen. Kim ist sich durchaus darüber im Klaren, dass diese Sichtweise voraussichtlich nicht überall auf Gegenliebe stoßen wird, sieht darin aber nichtsdestotrotz eine Notwendigkeit: „This will cause a chill in those physicalists who want to eat the cake and have it too – that is, those who want both the irreducibility and causal efficacy of the mental. I believe that the question no longer is whether or not those of us who want to protect mental causation find mind-body reductionism palatable. What has become increasingly clear (...) is that if we want robust mental causation, we had better be prepared to take reductionism seriously, whether we like it or not.“²⁰⁴

An späterer Stelle legt Kim nochmals ausführlicher sein Exklusionsargument dar. Zu diesem Zweck stellt er zunächst den Adressaten des Arguments vor – den nicht-reduktiven Physikalismus. Eine präzise Definition dieses Begriffs anzugeben sei, nach Kim, äußerst schwierig,

²⁰³ Vgl. Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 20f.

²⁰⁴ Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 22.

weil es nämlich keinen allgemeinen Konsens darüber gebe, was erstens ‚Physikalismus‘ genau bedeutet und wie zweitens, ‚Reduktion‘ verstanden werden solle. Allerdings dürften, Kim zufolge, die meisten Anhänger eines nicht-reduktiven Physikalismus den folgenden drei Prinzipien zustimmen:

In any case, most will agree that the following three doctrines are central to nonreductive physicalism: mind-body-supervenience, the physical irreducibility of the mental, and the causal efficaciousness of the mental. (...)

Supervenience. Mental properties strongly supervene on physical/biological properties. That is, if any system *s* instantiates a mental property *M* at *t*, there necessarily exists a physical property *P* such that *s* instantiates *P* at *t*, and necessarily anything instantiating *P* at any time instantiates *M* at that time.

(...)

Irreducibility. Mental properties are not reducible to, and are not identical with, physical properties.

(...)

Causal efficacy. Mental properties have causal efficacy – that is, their instantiations can, and do, cause other properties, both mental and physical to be instantiated.²⁰⁵

Nach Kims Verständnis ist Supervenienz allem voran eine ontologische These, die eine Art Abhängigkeit mit sich bringt. Diese Abhängigkeit rechtfertigt es zu sagen, dass eine mentale Eigenschaft in einem Organismus instanziiert ist, *weil* oder *aufgrund der Tatsache, dass* gleichzeitig in demselben Organismus eine bestimmte physikalische Eigenschaft (i.e. ihre Supervenienzbasis) realisiert ist. In diesem Sinne besteht nach der Supervenienzthese eine fundamentale Abhängigkeit des Mentalen vom Physikalischen. Offenkundig bedeutet Supervenienz in diesem Sinne weit mehr als nur das gleichzeitige gemeinsame Auftreten von physikalischen und mentalen Eigenschaften.²⁰⁶ Obwohl keine allgemein anerkannte Definition von ‚Reduktion‘ vorliegt, reicht nach Kim für das Exklusionsargument ein Verständnis in der Tradition von J.J.C. Smart²⁰⁷ aus: „(...) physically irreducible properties remain outside the physical domain – that is, if anything is physically reduced, it must be identical with some physical item. (...) If Xs are reduced to Ys, then Xs are nothing over and above the Ys.“²⁰⁸ Wenn also Eigenschaften physikalisch nicht reduzierbar sind, dann ist das gleichbedeutend damit, dass diese Eigenschaften nicht mit physikalischen Eigenschaften identisch sind. Die dritte von Kim herausgearbeitete Kernthese des nicht-reduktiven Physikalismus bezieht sich nun auf den kausalen Status der als irreduzibel postulierten mentalen Eigenschaften. Gerade Vertreter des nicht-reduktiven Physikalismus dürfte mit ihrer Theorie daran gelegen sein, dass mentale Eigenschaften eine Art Sonderstatus innehaben, der ihre kausale Wirksamkeit garantiert. Würden mentale Eigenschaften nämlich von vornherein als als kausal wirkungslos ein-

²⁰⁵ Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 33ff.

²⁰⁶ Vgl. Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 34.

²⁰⁷ Vgl. Smart, J.J.C.: „*Sensations and Brain Processes*“ in: *Philosophical Review*, Vol. 68, 1959, S. 141-156.

²⁰⁸ Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 34.

gestuft, so wäre die Frage, ob sie auf physikalische Eigenschaften reduziert werden können, ohnehin hinfällig. Schließlich könnten mentale Eigenschaften dann in keiner Kausalerklärung auftreten und es würde schlichtweg keine Rolle mehr spielen, ob sie vollständig auf physikalische Eigenschaften reduziert werden können oder nicht, weil die ‚Hauptbedrohung‘ einer Reduktion sicherlich in der kausalen Wirkungslosigkeit mentaler Eigenschaften (in ihrer Funktion als mentale Eigenschaften) zu sehen ist. Kurzum die Frage nach Reduzierbarkeit mentaler Eigenschaften wäre gänzlich uninteressant, wenn jene kausal wirkungslos wären. Andererseits könnte aber auch derart argumentiert werden: da physikalische Eigenschaften kausal wirksam sind, können kausal nicht wirksame mentale Eigenschaften sicherlich nicht auf diese reduziert werden, somit wäre die Irreduzibilitätsthese auf triviale Weise wahr. Es ist offensichtlich, dass ein nicht-reduktiver Physikalist sich auch dieser Argumentationsweise nicht anschließen wird. Schließlich wird er an der kausalen Wirksamkeit mentaler Eigenschaften großes Interesse hegen. Insofern besteht also ein Zusammenhang zwischen dem Irreduzibilitätsprinzip und dem Prinzip der kausalen Wirksamkeit mentaler Eigenschaften.²⁰⁹

Die treibende Kraft hinter dem Exklusionsargument nennt Kim selbst *Edwards's dictum* nach dem amerikanischen Philosophen Jonathan Edwards und definiert es wie folgt: „There is a tension between “vertical” determination and “horizontal” causation. In fact, vertical determination excludes horizontal causation.“²¹⁰ Was er unter vertikaler Determination versteht, illustriert Kim an folgendem Beispiel: Ein Bronzeklumpen hat zu irgendeinem Zeitpunkt bestimmte intrinsische Eigenschaften – z.B. Farbe, Form, Textur, Dichte, Härtegrad, elektronische Leitfähigkeit, usw. Die meisten würden der Aussage zustimmen, dass der Bronzeklumpen diese Eigenschaften aufgrund einer bestimmten Mikrostruktur hat – d.h. er ist aus verschiedenen Molekülen in einer bestimmten Anordnung zusammengesetzt. Die Makroeigenschaften des Bronzeklumpens sind somit vertikal determiniert durch seine aktuelle Mikrostruktur. Der Begriff ‚vertikal‘ rührt daher, weil die Mikro-Makro-Dimension für gewöhnlich auf der Y-Achse eines Koordinationsystems abgebildet wird. Auf der X-Achse hingegen verläuft der Zeitpfeil. Daraus leitet sich auch die ‚horizontale‘ Verursachung ab – von einem kausalen Standpunkt aus betrachtet, hat der Bronzeklumpen dann zu einem Zeitpunkt t die Eigenschaften, die er hat, weil er zu einem vorherigen Zeitpunkt $t - \Delta t$ die Eigenschaften hatte, die er zu diesem Zeitpunkt hatte. Somit determiniert die Vergangenheit die Zukunft und die Zukunft hängt von der Vergangenheit ab. Hieraus schlussfolgert Kim: „So we have here

²⁰⁹ Vgl. Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 35.

²¹⁰ Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 36.

two purported determinative relationships orthogonal to each other: vertical micro-macro mereological and horizontal past-to-future causal determination.²¹¹

Am Bronzeklumpen-Beispiel verdeutlicht sieht das folgendermaßen aus. Auf die Frage ‚Warum ist der Bronzeklumpen zum Zeitpunkt t gelb?‘ gibt es zwei mögliche Antworten. Die erste Möglichkeit besteht darin zu sagen, dass der Bronzeklumpen gelb ist, weil seine Oberfläche die mikrostrukturelle Eigenschaft M zum Zeitpunkt t hat. Die zweite mögliche Antwort könnte lauten, der Bronzeklumpen ist gelb, weil er zum Zeitpunkt $t - \Delta t$ gelb war. Um die Stärke des Exklusionsarguments begreifen zu können, muss man die Spannung zwischen diesen beiden Erklärungen erfassen. Denn für die erste Erklärung spielt es keine Rolle, was vor dem Zeitpunkt t passierte – worauf es einzig und allein ankommt ist, dass die Oberfläche des Klumpens die mikrostrukturelle Eigenschaft M hat. Wenn die Oberfläche des Klumpens zum Zeitpunkt t hingegen nicht die Mikroeigenschaft M (oder eine vergleichbare Mikroeigenschaft, die dieselben Wellenlängen des Lichts absorbiert) hat, könnte der Bronzeklumpen auch nicht gelb sein. Bei dieser Erklärung ist alles, was *zeitlich vor* t passierte völlig irrelevant für die gelbe Farbe des Bronzeklumpens. Das Haben der Mikroeigenschaft M zum Zeitpunkt t ist *allein* hinreichend für den Umstand, dass der Bronzeklumpen zum Zeitpunkt t gelb ist.²¹²

Edwards selbst vertrat die Position, dass es keine über die Zeit hinweg existierenden Objekte gebe, vielmehr sei Gott zu jedem Zeitpunkt die erhaltende Ursache seiner Schöpfung, der Welt. In jedem Moment erschaffe Gott also die Welt aufs Neue aus dem Nichts und erhalte sie kontinuierlich am Sein. Demnach ist Gott die vertikale Determination der Welt, die jegliche horizontale Verursachung ausschließt. Auf die Beziehung zwischen Mentalem und Physikalischem übertragen, bedeutet das:

It is simple to see how Edwards’s dictum applies to the mind-body case, causing trouble for mental causation. Mind-body supervenience, or the idea that the mental is physically “realized” – in fact, any serious doctrine of mind-body dependence will do – plays the role of vertical determination or dependence, and mental causation, or any “higher-level” causation, is the horizontal causation at issue. The tension between vertical determination and horizontal causation, or the former’s threat to preempt and void the latter, has been, at least for me, at the heart of the worries about mental causation.²¹³

Nachdem Kim nun die Bühne bereitet hat, präsentiert er eine noch weiter ausgearbeitete Version des Exklusionsarguments in Argumentform, die in zwei Stadien abläuft, wobei Kim für

²¹¹ Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 36.

²¹² Vgl. Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 37.

²¹³ Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 38.

das zweite Stadium zwei unterschiedliche Möglichkeiten präsentiert das Argument zu vervollständigen.²¹⁴

Stadium 1

Annahme: Es gibt Fälle von Mental-zu-Mental Verursachung, wobei M und M* mentale Eigenschaften sind.

(1) M verursacht M*.

Das Exklusionsargument geht von der Annahme aus, dass es – zumindest manchmal – Begebenheiten gibt, bei denen mentale Ursachen mentale Wirkungen haben – oder in anderen Worten bei denen mentale Ereignisse andere mentale Ereignisse verursachen. Es ist zu vermuten, dass die Wahrheit dieser Annahme zunächst von jedem Menschen anerkannt wird, jedenfalls solange diese Person nach ihrem gesunden Menschenverstand urteilt. Natürlich gibt es philosophische (Extrem-)Positionen, die selbst die Hypothese von der Existenz von Geist-zu-Geist-Verursachung bestreiten würden, aber es ist anzunehmen, dass auch die Verfechter solcher Theorien in ihrem Alltagsverständnis mutmaßlich von der Richtigkeit der These ausgehen.²¹⁵ Nehmen wir also mit Kim an, dass M M* verursacht, wobei M und M* hierbei als mentale Eigenschaften zu verstehen sind. Allerdings treten Eigenschaften als solche natürlich nicht in Kausalrelationen zueinander. Die Redeweise ‚M verursacht M*‘ steht verkürzt für ‚ein Vorkommen von M zum Zeitpunkt *t* verursacht ein Vorkommen von M* zum Zeitpunkt *t*‘ bzw. ‚eine Instanziierung von M zum Zeitpunkt *t* verursacht eine Instanziierung von M* zum Zeitpunkt *t*‘.

Aus der Supervenienzthese folgt die zweite Prämisse des Arguments:

(2) Für irgendeine physikalische Eigenschaft P* gilt: M* hat P* als seine Supervenienzbasis.

²¹⁴ Für die gesamte Darstellung des Exklusionsarguments, vgl. Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 39-45. Für eine bessere Lesbarkeit des Textes verzichte ich darauf an jedem Absatzende einen Verweis auf Kims Darstellung des Exklusionsargument einzufügen.

²¹⁵ Da Theorien ein archetypisches Beispiel für mentale Entitäten sind, würde sich auch eine Theorie T, die die Existenz mentaler Entitäten bestreitet, selbst den Boden entziehen. Denn wäre die Theorie T wahr, i.e. gäbe es keine mentalen Entitäten, dann könnte es überhaupt keine Theorien geben, also würde auch die Theorie T nicht existieren und könnte somit auch nicht wahr sein. (Vgl. z.B. Beckermann, Ansgar: *Analytische Einführung in die Philosophie des Geistes*, 2. überarb. Auflage, Berlin, de Gruyter, 2001, S.262ff.

Aus (1) und (2) ergibt sich ein Spannungsverhältnis, wenn man die Frage stellt, warum M^* zu diesem Zeitpunkt instanziiert wurde oder welche Tatsache sich dafür verantwortlich zeichnet oder erklärt, dass M^* in diesem Moment eintritt. Denn wie es scheint gibt es auf diese Fragen zwei unterschiedliche, sich gegenseitig jeweils ausschließende Antworten – und zwar: ‚ M verursachte, dass M^* zu diesem Zeitpunkt instanziiert wurde‘ und ‚ P^* , die Supervenienzbasis von M^* , verursachte, dass M^* in diesem Moment instanziiert wurde‘. An dieser Stelle kommt Edwards’s dictum ins Spiel: Angenommen P^* findet statt, dann wird auch M^* eintreten – unabhängig davon, was vorher passiert ist, denn als Supervenienzbasis von M^* ist die Instanzierung von P^* zum Zeitpunkt t allein und für sich schon hinreichend für das Eintreten von M^* zum Zeitpunkt t . Dies würde selbst dann gelten, wenn M , die vermeintliche Ursache von M^* , nicht stattgefunden hätte – es sei denn das Eintreten von M hatte etwas mit dem Eintreten von P^* zu tun. Unmittelbar aus diesem Gedanken leitet Kim die dritte Prämisse seines Arguments ab:

(3) M verursacht M^* *indem* es die Supervenienzbasis P^* verursacht.

Mit dieser Annahme ist das erste Stadium des Arguments abgeschlossen. Bis zu diesem Punkt hat das Argument gezeigt, dass sofern mentale Eigenschaften auf physikalischen supervenieren, Geist-zu-Geist-Verursachung bereits Geist-zu-Körper-Verursachung einschließt. Erweist sich die Supervenienzthese also als wahr, dann ist es unmöglich, dass es Verursachung im mentalen Bereich gibt, ohne dass die Verursachung dabei auch gleichzeitig die Grenze zum ‚Reich‘ der Physik überschreitet. Dieser Befund ist von großer Bedeutung, denn er hat zur Folge, dass bei der Annahme einer Mikro-Makro-Supervenienz Kausalbeziehungen nicht auf einen Bereich beschränkt werden können – jede Verursachung auf einer Stufe S (die höher als die Basisstufe ist) zieht eine stufenübergreifende Verursachung von S zu $S - 1$ mit sich. Oder, um es mit Kims Worten auszudrücken: „In short, *level-bound causal autonomy is inconsistent with supervenience or dependence between the levels.*“²¹⁶ Außerdem zeigt das Exklusionsargument bereits an dieser Stelle, dass Geist-zu-Geist-Verursachung unter physikalistischen Vorzeichen genauso viele Probleme bereitet wie Geist-zu-Körper-Verursachung. Gemeinhin wird es so dargestellt als sei mentale Verursachung insbesondere im Hinblick auf die Frage, wie das Mentale im kausal geschlossenen Bereich der Physik kausal wirksam sein kann, problematisch. Allerdings zeigt sich, dass der Physikalismus alle Formen bzw. Richtungen men-

²¹⁶ Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 40.

taler Verursachung, eben auch Geist-zu-Geist-Verursachung, bedroht. Erst im zweiten Stadium des Exklusionsarguments wird auf Geist-zu-Körper-Verursachung Bezug genommen. Es gilt zu beachten, dass das Exklusionsargument im ersten Stadium keinerlei metaphysische Prinzipien hinsichtlich der Physik, wie etwa deren kausale Geschlossenheit, voraussetzt. Geist-Körper-Supervenienz ist die einzige fundamentale Prämisse, die präsupponiert wird.

Stadium 2

Im zweiten Stadium des Exklusionsarguments gibt es zwei mögliche Arten das Argument zu vervollständigen. Im Folgenden wird zunächst die Originalversion des Arguments und später eine neue Variation vorgestellt.

Erster möglicher Abschluss des Arguments:

Lenkt man nun die Aufmerksamkeit auf M, die vermeintliche Ursache von M*, folgt aus der Supervenienzzannahme die vierte Prämisse Exklusionsarguments:

(4) M hat eine physikalische Supervenienzbasis, und zwar P.

Nun gibt es gute Gründe anzunehmen, dass P die Ursache von P* ist. Denn P ist (zumindest) nomologisch hinreichend für M und das Eintreten von M hängt von dem Eintreten von P in dieser Situation ab bzw. ist determiniert durch das Stattfinden von P. Allerdings gilt per Annahme, dass M die Ursache von P* ist. Es scheint jedoch so, als würde auch P die adäquaten Voraussetzungen dafür mitbringen die Ursache von P* zu sein. Daraus ergibt sich folgende Aussage als fünfter Schritt des Exklusionsarguments:

(5) M verursacht P* und P verursacht P*.

Hierbei kann P's Verursachung von P* nicht als eine Kausalkette mit der Zwischenstation M betrachtet werden. Ein Grund hierfür ist, dass es sich bei der P-zu-M-Relation nicht um eine Kausalbeziehung handelt. Ein zweiter Grund ist, dass beide exakt zum selben Zeitpunkt stattfinden, weil M auf P superveniert. Somit können P und M nicht in einem Ursache-Wirkungs-Verhältnis zueinander stehen, zumal sich (nach dem Allgemeinverständnis von Kausalität) Ursachen immer zeitlich *vor* ihren Wirkungen ereignen müssen. Zudem wird M später durch

das Einführen des Exklusionsprinzips und des Prinzips der kausalen Abgeschlossenheit der Physik als Ursache von P^* ausgeschlossen.

Die nächste Prämisse des Arguments ist das Irreduzibilitätsprinzip:

(6) $M \neq P$.

Es ist offensichtlich, dass auch zwischen den Prämissen (5) und (6) ein metaphysisches Spannungsverhältnis besteht. Denn die Darstellung des Arguments bis dato hätte zur Folge, dass P^* zwei unterschiedliche Ursachen hat, die jede für sich alleine genommen schon hinreichend für das Eintreten von P^* ist. Evidenterweise kommt an dieser Stelle das kausale Exklusionsprinzip ins Spiel. Eine Formulierung dieses Prinzips könnte wie folgt lauten:

Exklusionsprinzip. Ein einzelnes Ereignis kann zu einem gegebenen Zeitpunkt nicht mehr als eine hinreichende Ursache haben – es sei denn, es handelt sich um einen Fall genuiner kausaler Überdeterminierung.

Prämisse (7) ist schlichtweg die Annahme, dass P^* nicht kausal überdeterminiert ist.

(7) P^* ist nicht kausal überdeterminiert durch M und P .

Aus dem Exklusionsprinzip folgt, dass entweder M oder P als Ursache von P^* eliminiert werden muss. Die Konklusion (8) des Arguments folgt nun aus einer weiteren metaphysischen Annahme – der Annahme der kausalen Geschlossenheit der Physik, kurzum dem Geschlossenheitsprinzip.

Geschlossenheitsprinzip. Wenn ein physikalisches Ereignis eine Ursache hat, die zum Zeitpunkt t eintritt, dann hat es eine physikalische Ursache, die zum Zeitpunkt t eintritt.

Das Geschlossenheitsprinzip dient nun als Heuristik für die Entscheidung welche der beiden vermeintlichen Ursachen aus der Kausalerklärung auszuschließen ist. Es ist leicht ersichtlich, dass es demnach die mentale Ursache M ist, die aufgegeben werden muss. Schließlich ist P^* per definitionem ein physikalisches Ereignis – daher muss P^* nach dem Geschlossenheitsprinzip ein anderes physikalisches Ereignis als hinreichende Ursache haben (natürlich nur

unter der Voraussetzung, dass P^* überhaupt irgendeine hinreichende Ursache hat). Es ist manifest, dass dieses andere physikalische Ereignis P sein muss. Würde man sich hingegen für M als Ursache von P^* entscheiden, so käme wieder das Geschlossenheitsprinzip zum Zuge. Demzufolge bräuchte P^* eine physikalische Ursache, etwa P_1 (was wiederum wohl nichts anderes als P sein könnte). Danach sähe man sich durch das Exklusionsprinzip erneut gezwungen eine Entscheidung zwischen M und P_1 (d.h. P) als Ursache von P^* zu fällen. Solange man also nicht für P als Ursache von P^* votiert, scheint ein infinites Regress zu drohen, bei dem man fortwährend vor der Entscheidung stünde: Ist M oder irgendeine physikalische Ursache P_x die eigentliche Ursache von P^* ? Somit gelangt man schließlich zur Schlussfolgerung des Exklusionsarguments:

- (8) Die vermeintliche mentale Ursache M wird von der physikalischen Ursache P ausgeschlossen. Das heißt P und nicht M ist die Ursache von P^* .

Es ist interessant zu eruieren wie das Exklusions- und das Geschlossenheitsprinzip gemeinsam zur der epiphänomenalen Konklusion (8) führen. Das Exklusionsprinzip verhält sich nämlich neutral hinsichtlich des Wettbewerbs zwischen M und P – es postuliert lediglich, dass entweder die physikalische Ursache P oder die mentale Ursache M aus der Kausalerklärung ausgemustert werden muss. Das Exklusionsprinzip allein gibt also noch nicht an, *welche* der beiden Ursachen zu favorisieren ist. Diese asymmetrische Komponente tritt erst mit dem Geschlossenheitsprinzip zutage. Denn erst die kausale Geschlossenheit der Physik schließt die mentale Ursache aus. Wären die Vorzeichen umgekehrt, d.h. wäre etwa der mentale Bereich kausal geschlossen und der physikalische nicht, dann würde auch die mentale Ursache über die physikalische Ursache obsiegen. In einer solchen Form des Idealismus würde man dann auch nicht mehr von Problemen mentaler Verursachung sprechen, sondern wäre stattdessen mit Problemen physikalischer Verursachung konfrontiert.

Zweiter möglicher Abschluss des Arguments:

Neben der gerade ausführlich geschilderten Originalversion des Exklusionsproblems, präsentiert Kim auch eine andere Art und Weise das zweite Stadium des Exklusionsarguments zu komplettieren. Diese Methode zieht aus der dritten Prämisse (zur Erinnerung: (3) M verursacht M^* *indem* es die Supervenienzbasis P^* verursacht.) folgenden Schluss:

(4) M ist die Ursache von P*.

Aus dem Geschlossenheitsprinzip lässt sich die nächste Prämisse ableiten:

(5) P* hat eine physikalische Ursache – sagen wir P – die zum selben Zeitpunkt wie M eintritt.

Die sechste Prämisse ist – genau wie in der Originalversion des Exklusionsarguments – die Irreduzibilitätsthese:

(6) $M \neq P$.

Woraus sich wiederum direkt die nächste Prämisse ableiten lässt:

(7) Daher hat P* zwei unterschiedliche Ursachen, und zwar M und P. Zudem ist dies kein Fall einer kausalen Überdeterminierung.

Infolge des Exklusionsprinzips muss eine Entscheidung zwischen M und P gefällt werden.

(8) Entweder muss man M oder P fallen lassen.

Diese kann nach dem Geschlossenheitsprinzip und dem Exklusionsprinzip nur folgendermaßen lauten:

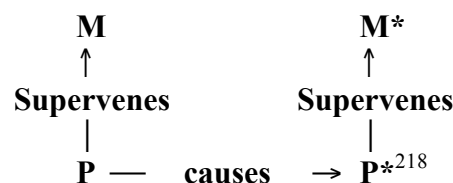
(9) M muss fallen gelassen werden, P bleibt erhalten.

Das zweite ‚Finale‘ des Exklusionsarguments ist augenscheinlich einfacher als die Originalversion, denn hier wird die Supervenienzthese nicht als eine Prämisse des Arguments benötigt. Infolgedessen wird auch die Behauptung, M’s Supervenienzbasis P habe einen gültigen Anspruch darauf die Ursache von P* zu sein, umgangen.

Andererseits ist die Originalversion in einigen Gesichtspunkten intuitiv einleuchtender als die modifizierte Variante: so wird in ihr Edwards’s dictum expliziter. Zudem zeigt sie auf herausragende Weise wie vermeintliche höherstufige Kausalrelationen von kausalen Prozessen einer niedrigeren Stufe außer Kraft gesetzt bzw. überflüssig gemacht werden. Jedenfalls dient das

zweite Stadium des Exklusionsarguments vornehmlich dazu die Schwächen von ‚Abwärts-Verursachung‘ aufzuzeigen – und zwar: „(...) that *the assumptions of causal exclusion and lower-level causal closure disallow downward causation*.“²¹⁷

Das Ergebnis der Originalversion des Exklusionsarguments stellt Kim mit diesem Schema dar:



Nach dieser Darstellung besteht ausschließlich zwischen P und P* eine Kausalrelation, dementsprechend stehen M und M* nicht – wie zu Beginn des Exklusionsarguments angenommen (vgl. Prämisse (1)) – in einer Ursache-Wirkungs-Beziehung zueinander. Die Verursachung spielt sich also rein auf der physikalischen Ebene ab. Die Kausalbeziehung zwischen den beiden mentalen Eigenschaften M und M* kann eliminiert werden, weil sie jeweils auf gewissen physikalischen Eigenschaften (nämlich P und P*) supervenieren, die bereits in einer Verursachungsbeziehung zueinander stehen. Daher leisten auch weder M noch M* einen Beitrag dazu, die kausale Struktur mitzugestalten – sie supervenieren lediglich auf den physikalischen Eigenschaften, die die Kausalstruktur bereits vollkommen festlegen. Durch die Supervenienzbeziehung zwischen M und P (respektive M* und P*) kann zusammen mit der Kausalbeziehung zwischen P und P* eine kontrafaktische Abhängigkeit zwischen den beiden mentalen Eigenschaften und zwischen den mentalen und den physikalischen Eigenschaften postulieren. In einer solchen kontrafaktischen Dependenz ist jedoch sicherlich keine Kausalbeziehung im eigentlichen Wortsinn zu sehen.

Die zweite Vervollständigung des Exklusionsarguments liefert ein etwas weniger striktes Bild: der senkrechte Supervenienz-Pfeil von P zu M entfällt. Gewiss muss M auch in der zweiten Version des Exklusionsarguments eine Supervenienzbasis besitzen – allerdings muss diese nicht mehr unbedingt, wie in der Originalversion, die Ursache von P* sein. Das Ergebnis bleibt jedoch dasselbe: die vermeintlichen Kausalbeziehungen zwischen M und M*, sowie zwischen M und P* müssen der zugrundeliegenden Kausalbeziehung zwischen P und P* den Vorrang einräumen. Somit werden die mentalen Eigenschaften M und M* durch die physikalische Eigenschaft P aus der Kausalerklärung von P* ausgeschlossen.

²¹⁷ Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 44.

²¹⁸ Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005, S. 45.

Kritik am Exklusionsargument

Überdeterminierung I

Eine sehr umfangreiche Kritik an Kims Exklusionsargument stellen Thomas Crisp und Ted Warfield in ihrem 2001 in der Zeitschrift *Noûs* erschienenen Artikel *Kim's Master Argument* vor. Dazu fassen sie zunächst Kims Argument zusammen: Crisp und Warfield zufolge bedient sich das Exklusionsargument dreier metaphysischer Prinzipien – und zwar denen der Supervenienz, der Geschlossenheit der Physik und des Eigenschaftsdualismus. Diese drei Prinzipien werden von Crisp und Warfield wie folgt definiert und stehen in folgendem Zusammenhang:

Supervenience – Necessarily, for any mental property M, if anything has M, there exists a physical base property P such that it has P, and, necessarily, anything that has P has M.

(...)

Closure – Every caused physical event has a physical cause.

(...)

Property Dualism – Mental properties are distinct from physical properties.

(...)

In broad outline, Kim's supervenience argument goes as follows:

P1. Either Supervenience holds or it does not.

P2. If it fails to hold, then, if Closure and Property Dualism, mental causation is 'unintelligible'.

P3. If it holds, then, if Closure and Property Dualism, mental causation is 'unintelligible'.

C1. Therefore, if Closure and Property Dualism, then mental causation is 'unintelligible'.²¹⁹

Bei der Diskussion von Kims Exklusionsargument stellen Crisp und Warfield zunächst fest, dass es sich bei dem Argument um einen formal gültigen Schluss handelt – d.h. das Argument ist von ‚handwerklich-technischer‘, oder besser von logischer Seite betrachtet, einwandfrei. Aus diesem Grund müsse man für eine qualifizierte Kritik des Arguments die einzelnen Prämissen und der Schluss inhaltlich genauestens durchleuchten.

Es ist augenscheinlich, dass P1 logisch betrachtet eine Tautologie ist und daher wahr sein muss – die erste Prämisse ist eine simple Anwendung des *tertium non datur* auf das Supervenienzprinzip: eine der beiden Aussagen ‚Supervenienz trifft zu‘ und ‚Supervenienz trifft nicht zu‘ muss schlichtweg wahr sein – dies ist über jeden Zweifel erhaben. Daher wenden sich Crisp und Warfield postwendend der zweiten Prämisse des Arguments zu. Diese besagt, dass wenn Supervenienz nicht zutrifft, gleichzeitig aber der Bereich der Physik kausal geschlossen und der Eigenschaftsdualismus wahr ist, mentale Verursachung unverständlich wird, oder besser gesagt, nicht mehr kohärent denkbar ist. Das Problem besteht dann nämlich darin zu

²¹⁹ Crisp, Thomas & Warfield, Ted: „Kim's Master Argument“, in: *Noûs*, Vol. 35:2, 2001, S. 305f.

erklären, wie das Mentale im kausal abgeschlossenen Bereich der Physik kausal wirksam werden kann, wenn Mentales weder auf Physikalischem superveniert, noch mit ihm identisch ist. Wenn es also keine Geist-Gehirn-Supervenienz gibt, verschwindet somit auch die Intelligibilität mentaler Verursachung. Nach Kim besteht der offensichtlichste Grund für die Wahrheit dieser Behauptung in der Wahrheit der kausalen Geschlossenheit der Physik. Der zugrundeliegende Gedanke scheint der zu sein, dass wenn der Eigenschaftsdualismus zutrifft und Geist-Gehirn-Supervenienz nicht, Geist-Körper-Verursachung die Falschheit der kausalen Geschlossenheit der Physik beinhaltet. Aus dieser Idee leitet Kim direkt die Prämisse P2 ab: wenn die Existenz von Geist-zu-Körper-Verursachung zusammen mit dem Fehlen von Supervenienz *logisch inkonsistent* mit der Wahrheit der kausalen Geschlossenheit der Physik und des Eigenschaftsdualismus ist, dann besteht für Vertreter dieser beiden Hypothesen kein Hoffnungsschimmer zu verstehen wie eine solche Verursachung überhaupt möglich sein könnte. An dieser Stelle wenden Crisp und Warfield ein, dass Anhänger des Eigenschaftsdualismus und des Prinzips der kausalen Geschlossenheit, selbst wenn sie annehmen, dass Supervenienz nicht zutrifft, nicht, wie von Kim behauptet, verpflichtet sind mentale Verursachung als *gänzlich* nicht-intelligibel abzustempeln. Vielmehr sei davon nur mentale Verursachung in Geist-Körper-Richtung betroffen, denn die Intelligibilität von Geist-zu-Geist-Verursachung bleibe davon unberührt. Aber auch was Geist-zu-Körper-Verursachung betrifft, sind Crisp und Warfield Kims Ausführungen gegenüber skeptisch, weil diese nämlich allem Anschein nach auf einer falschen Annahme fußen. Nach Crisp und Warfield hat die Existenz von Geist-zu-Körper-Verursachung gemeinsam mit dem Fehlen von Supervenienz und der Wahrheit des Eigenschaftsdualismus *nicht*, wie von Kim präsupponiert, die Falschheit des Geschlossenheitsprinzips zur Folge.²²⁰ Zur Untermauerung ihrer These führen Crisp und Warfield diesen Gedankengang an:

Let us say that mental to physical causation occurs when an event E1 causes a physical event E2 in virtue of E1's having a mental property. Now suppose the world would be such that every instance of a certain mental property M is accompanied by an instance of some physical property or other, but that M doesn't supervene on any of these physical properties: each is such that it is nomologically possible for an event to instantiate it without also instantiating M. Suppose further that M-events cause P*-events in virtue of their having M. Kim must claim that these assumptions imply the falsity of Closure, but he is mistaken. For we've supposed (and this supposition is consistent with denying Supervenience) that every instance of M occurs together with an instance of some physical property P. And we see no reason for thinking that it couldn't *also* be the case that there is a causal law for each P to the effect that the instantiation of P is causally sufficient for the instantiation of P*. Thus we see no reason for thinking that the conjunction of the denial of Supervenience, Property Dualism, and the existence of mental to physical causation implies a violation of Closure.²²¹

²²⁰ Vgl. Crisp, Thomas & Warfield, Ted: „Kim's Master Argument“, in: *Noûs*, Vol. 35:2, 2001, S. 306f.

²²¹ Crisp, Thomas & Warfield, Ted: „Kim's Master Argument“, in: *Noûs*, Vol. 35:2, 2001, S. 307.

In diesem von Crisp und Warfield geschilderten Szenario, supervenieren mentale Eigenschaften nicht auf physikalischen Eigenschaften (\neg Supervenienz), mentale Eigenschaften unterscheiden sich von physikalischen Eigenschaften (Eigenschaftsdualismus) und es findet Geist-zu-Körper-Verursachung statt (mentale Verursachung). Dennoch führt die Konjunktion dieser Annahmen, so Crisp und Warfield, nicht zwangsläufig zur Falschheit des Geschlossenheitsprinzips. Denn es lässt sich widerspruchsfrei annehmen, dass M-Ereignisse P*-Ereignisse aufgrund ihres Habens von M verursachen *und zugleich* ist auch ein Kausalgesetz für jedes P vorstellbar, dass Instanziierungen von P als (kausal) hinreichend für Instanziierungen von P* annimmt. An dieser Stelle könnte man einwenden, dass Crisp und Warfield Kims These von der kausalen Geschlossenheit missverstehen und deshalb auch sein gesamtes Argument falsch darstellen. In seinem Buch *Mind in a Physical World* präsentiert Kim nämlich eine Definition des Prinzips der kausalen Geschlossenheit, die von Crisps und Warfields Darstellung abweicht: „One way of stating the principle of physical causal closure is this: If you pick any physical event and trace out its causal ancestry or posterity, that will never take you outside the physical domain. That is, no causal chain will ever cross the boundary between the physical and the nonphysical.“²²² Es ist offensichtlich, dass diese Definition eine weitaus stärkere Annahme als Crisps und Warfields Definition des Geschlossenheitsprinzips ist. Dementsprechend taufen Crisp und Warfield diese These ‚Exclusion‘ und definieren sie folgendermaßen:

Exclusion – Every caused physical event has only physical causes.²²³

Offenkundigerweise verliert mentale Verursachung (in Geist-Körper-Richtung) tatsächlich ihre Intelligibilität, wenn Kims Prinzip der kausalen Geschlossenheit dem von Crisp und Warfield vorgestellten ‚Exclusion‘ und nicht dem von ihnen präsentierten ‚Closure‘ entspricht. Denn Geist-zu-Körper-Verursachung findet definitionsgemäß dann statt, wenn ein Ereignis E1 ein physikalisches Ereignis E2 deshalb verursacht, weil E1 gewisse mentale Eigenschaften hat. Wenn aber physikalische Ereignisse *allein* aufgrund ihrer physikalischen Eigenschaften in Kausalbeziehungen treten, dann kann es keine mentale Verursachung in diesem Sinne geben. Falls Kims Definition des Prinzips der kausalen Geschlossenheit also mit Crisps und Warfields ‚Exclusion‘ und nicht mit deren ‚Closure‘ übereinstimmt, dann verliert auch die Kritik von Crisp und Warfield an der zweiten Prämisse des Exklusionsarguments (P2 in ihrer eigenen Darstellung) entscheidend an Überzeugungskraft. Um genau dies abzuwenden, argumentieren Crisp und Warfield, dass Kim mit dem Geschlossenheitsprinzip nicht ‚Exclusion‘

²²² Kim, Jaegwon: *Mind in a Physical World – An Essay on the Mind-Body Problem and Mental Causation*, Cambridge (Mass.), MIT Press, 2000, S.40.

²²³ Crisp, Thomas & Warfield, Ted: „Kim’s Master Argument“, in: *Noûs*, Vol. 35:2, 2001, S. 307.

meinen *kann*. Schließlich scheint es so, als vertrete Kim die Ansicht, dass die Idee der Supervenienz nicht-reduktiven Physikalisten wenigstens einen Funken Hoffnung spenden sollte Geist-zu-Körper-Verursachung zu verstehen, ohne gleichzeitig die Identität von Mentalem und Physikalischen vorauszusetzen. Wird das Geschlossenheitsprinzip jedoch im Sinne von ‚Exclusion‘ verstanden, dann gibt es nicht einmal mehr diesen Silberstreifen am Horizont für nicht-reduktive Physikalisten. Zweifelsohne kann es trivialerweise keine Geist-zu-Körper-Verursachung geben, wenn Eigenschaftsdualismus und ‚Exclusion‘ gleichzeitig wahr sind. Wenn also die Supervenienzthese dazu dient mentale Verursachung in Geist-Körper-Richtung im nicht-reduktiven Physikalismus mit der kausalen Geschlossenheit der Physik zu versöhnen, dann kann unter Geschlossenheit nicht ‚Exclusion‘ verstanden werden. Die originäre Funktion der Supervenienzthese scheint aber tatsächlich eben genau darin zu bestehen. Somit wäre es alles andere als fair, wenn man das Geschlossenheitsprinzip im Sinne von ‚Exclusion‘ definieren würde. Auf diese Weise würde nämlich die Notwendigkeit der Identität zwischen Mentalem und Physikalischen gleich in die – vermeintlich neutrale – Ausgangsposition gepflanzt werden. In anderen Worten würde es sich bei einem solchen Vorgehen um eine klassische *petitio principii* handeln: die Konklusion des Arguments wäre bereits (in diesem Falle implizit) in den Prämissen enthalten.

Eine weitere Möglichkeit Kims Exklusionsargument anzugreifen besteht nach Crisp und Warfield darin Prämisse P3 anzugreifen. Zu diesem Zweck untersuchen Crisp und Warfield Kims Argumentation für P3 – diese vollzieht sich ihnen zufolge in zwei Schritten. Der erste Schritt ist der Versuch zu beweisen, dass die Existenz von Geist-zu-Geist-Verursachung schon die Existenz von Geist-zu-Körper-Verursachung mit einschließt. Der zweite Schritt ist es zu argumentieren, dass Geist-zu-Körper-Verursachung nicht-intelligibel ist, wenn man gleichzeitig von der Wahrheit des Geschlossenheits- und Supervenienzprinzips, sowie des Eigenschaftsdualismus ausgeht.

Für die erste Stufe des Arguments für P3 – also die Annahme, dass Geist-zu-Geist-Verursachung bereits Geist-zu-Körper-Verursachung bedingt – argumentiert Kim wie oben bereits ausführlich dargelegt: angenommen, die Instanziierung einer mentalen Eigenschaft M verursacht die Instanziierung einer anderen mentalen Eigenschaft M*. Aus der Supervenienzthese folgt, dass M* eine physikalische Supervenienzbasis P* benötigt. Nun scheint es auf die Frage was die Instanziierung von M* verursacht zwei Antworten zu geben – einerseits die mentale Eigenschaft M, andererseits M*s physikalische Supervenienzbasis P*. Zwischen diesen beiden Antworten besteht, laut Kim, ein starkes Spannungsverhältnis. Um einen Schiedsspruch in diesem Wettbewerb um die Erklärung für das Auftreten von M* zu erzielen,

führt Kim das Prinzip der kausalen bzw. explanatorischen Exklusion ein. Dieses besagt, dass es für ein Ereignis immer nur *eine* vollständige und unabhängige (kausale) Erklärung geben kann. Demzufolge ist anzunehmen, dass Kim M und P* für zwei vollständige und voneinander unabhängige Erklärungen für die Instanziierung von M* hält. Wenn dies aber der Fall wäre, dann würde dadurch aber das Prinzip der kausalen Exklusion verletzt. Der einzige Ausweg, den Kim in dieser verfahrenen Situation sieht, besteht in der Annahme, dass M M* verursacht *indem* es P* verursacht. Daraus lässt sich folgern, dass Geist-zu-Geist-Verursachung also nur dann möglich ist, wenn gleichzeitig auch Geist-zu-Körper-Verursachung möglich ist oder in anderen Worten die Existenz dieser bereits die Existenz jener impliziert.

Hinsichtlich dieses ersten Schrittes von Kims Argument für P3, machen Crisp und Warfield drei Einwendungen. Erstens ist es alles andere als klar, warum man das Prinzip der kausalen Exklusion verletzt sein sollte, wenn man annimmt, dass sowohl M als auch P* Erklärungen für M* sind. Schließlich müssten die Erklärungen dem besagten Prinzip zufolge *unabhängig* voneinander sein – M und P* sind jedoch nicht unabhängig voneinander. Denn gemäß des Supervenienzprinzips hat M eine physikalische Eigenschaft als Supervenienzbasis. Wenn nun aber M's Supervenienzbasis die Ursache von P* ist, dann scheint es so als wären M und P* nicht zwei voneinander *unabhängige* Kausalerklärungen für das Auftreten von M*. Zweitens ist unklar, wie diese beiden Erklärungen für M* im Wettbewerb zueinander stehen sollen. Schlussendlich handelt es sich dabei um fundamental unterschiedliche Arten von Erklärungen: das Sich-Ereignen von M ist eine *diachronische, kausale* Erklärung für das Eintreten von M*, wohingegen die Instanziierung von P* eine *synchrone, nicht-kausale* Erklärung für das Eintreten von M* darstellt (vgl. Edwards dictum). Im Allgemeinen ist die Art von Erklärung, die eine subveniente Basiseigenschaft für die Instanziierung einer supervenienten Eigenschaft liefert, eine nicht-kausale Erklärung. Da es sich hier um zwei grundverschiedene Arten von Erklärung handelt, ist nicht ersichtlich auf welche Art und Weise diese miteinander konkurrieren könnten. Um dies zu illustrieren führen Crisp und Warfield folgendes Beispiel an: die Eigenschaft ‚Masern zu haben‘ kann auf unterschiedliche Weise verstanden werden. Einerseits kann man diese Eigenschaft als höherstufige, funktionale Eigenschaft eines physikalischen Systems auffassen, die durch bestimmte physikalische Eigenschaften erster Ordnung des Systems realisiert ist. Andererseits stellt die Eigenschaft mit einem bestimmten Virus infiziert zu sein eine kausale Erklärung für das Haben von Masern dar. Auf die Frage warum jemand Masern hat, gibt es dementsprechend auch zwei angemessene Antworten. Und zwar, a.) weil er sich mit einem bestimmten Virus infiziert hat oder b.) weil diese Person gerade

bestimmte physikalische Basiseigenschaften instanziiert, auf denen die Eigenschaft ‚Masern haben‘ superveniert. Dies sind zwei Arten ein und dasselbe Phänomen zu beschreiben – die eine kausal und diachronisch, die andere nicht-kausal und synchron. Allerdings konkurrieren diese beiden Erklärungsarten nicht miteinander und jede für sich kann als eine adäquate Erklärung für das Haben von Masern gelten.²²⁴ Crisp und Warfield zufolge, hält Kim denjenigen, die bezweifeln, dass zwischen beiden Erklärungen ein Spannungsverhältnis besteht folgendes Prinzip entgegen: Um die Instanziierung einer supervenienten Eigenschaft zu verursachen, muss eine ihrer Basiseigenschaften verursacht werden. An dieser Stelle greift der dritte Kritikpunkt Crisps und Warfields hinsichtlich der ersten Stufe von Kims Argument für P3 – sie analysieren die Motivation Kims für dieses Prinzip. Für die Ausformulierung ihres Gedankengangs beginnen sie damit, zwei Lesarten des Prinzips darzulegen:

Here are two slightly different (nonequivalent) formulations of Kim’s principle:

- (R1) In order for X to cause a supervenient property to be instantiated, X must cause one of its base properties to be instantiated.
- (R2) In order for X to cause a supervenient property to be instantiated, the instantiation of one of its base properties must be caused.²²⁵

Für die Etablierung dieses Prinzips ersinnt Kim einige Beispiele, wobei seine bevorzugte Interpretation wohl R1 ist: Wenn man Kopfschmerzen hat, besteht der einzige Weg die Kopfschmerzen zu lindern darin, dass man kausal in die Hirnprozesse eingreift, auf denen die Kopfschmerzen supervenieren. Aber selbst wenn man, um die Kopfschmerzen abzustellen, die Gehirnprozesse auf denen die Schmerzen supervenieren ändern muss, stellt sich die Frage, ob wirklich ein und dieselbe Eigenschaft für *beides* kausal verantwortlich sein muss – also für das Verschwinden der Kopfschmerzen und für die Veränderung der ihnen zugrundeliegenden neuronalen Prozesse. Crisp und Warfield schlagen diesbezüglich eine andere Strategie vor:

Why not suppose that (a) what causes the disappearance of my headache is my instantiation of the property being-such-as-to-have-just-taken-an-aspirin, a higher-order functional property realized by certain of my first-order physical properties, and that (b) what causes the alternation of my headache’s underlying base properties is the instantiation of these first-order properties? This way of putting things looks to us to be coherent and so we find it far from obvious that Kim’s examples specifically motivate (R1) over (R2). But as applied to the issue under discussion, (R2) is clearly consistent with there being mental to mental causation without mental to physical causation. Thus we do not think that Kim’s reasoning here provides the compelling independent justification he promised for thinking that M causes M* by causing P*. The first step of Kim’s argument for (P3), we conclude, is a shaky step.²²⁶

²²⁴ Vgl. Crisp, Thomas & Warfield, Ted: „Kim’s Master Argument“, in: *Noûs*, Vol. 35:2, 2001, S. 307ff.

²²⁵ Crisp, Thomas & Warfield, Ted: „Kim’s Master Argument“, in: *Noûs*, Vol. 35:2, 2001, S. 311.

²²⁶ Crisp, Thomas & Warfield, Ted: „Kim’s Master Argument“, in: *Noûs*, Vol. 35:2, 2001, S. 311.

Obwohl Crisp und Warfield davon ausgehen, dass mit ihrer Kritik an der ersten Stufe auch das gesamte Argument für (P3) zurückgewiesen werden kann, formulieren sie weitere Einwände gegen die zweite Stufe des Arguments. Die zweite Stufe ist, wie oben bereits expliziert, das Argument, dass Geist-zu-Körper-Verursachung nicht-intelligibel ist, wenn gleichzeitig die Prinzipien der Supervenienz, der kausalen Geschlossenheit und des Eigenschaftsdualismus zutreffen. Eine Rekonstruktion des Gedankengangs könnte nach Crisp und Warfield in etwa so aussehen: Angenommen Supervenienz trifft zu, dann muss M eine Supervenienzbasis haben – etwa P. Nun gilt es den kausalen Status von P hinsichtlich P* zu überprüfen. Dabei sollte sich herausstellen, dass es mindestens genauso gute Gründe gibt P als Ursache von P* zu betrachten wie M als Ursache von P* zu erachten. Denn egal ob Verursachung als Nomologisches-hinreichend-Sein oder als kontrafaktische Abhängigkeit verstanden wird, in beiden Fällen scheint es gerechtfertigt zu sein, zu behaupten P verursache P*. Denn P erweckt den Eindruck eine (mindestens) ebenso gute Ursache für P* zu sein wie M. Das wiederum hat nach Kim zur Folge, dass es nun einen Überschuss an Ursachen für P* gibt – es scheint so als wäre P* kausal überdeterminiert. Wenn dies allerdings der Fall wäre, dann wären nicht-reduktive Physikalisten Kim zufolge dazu verurteilt anzunehmen, dass *jeder* Fall mentaler Verursachung systematisch überdeterminiert ist. Schließlich benötigen mentale Eigenschaften immer physikalische Supervenienzbasen, mit diesen stünden sie dann aber sozusagen im kausalen Wettbewerb um die Verursachung der subvenienten Eigenschaften. Stellt man sich hingegen eine mögliche Welt vor, die im höchsten Maße der aktuellen Welt gleicht, in der aber P nicht eintritt und somit M letztlich als einzige Ursache von P* übrig bliebe, dann würde dadurch das Geschlossenheitsprinzip verletzt, so Kim. Denn in diesem Fall hätte P* ausschließlich nicht-physikalische Ursachen. Wobei Kim natürlich stillschweigend voraussetzt, dass das Geschlossenheitsprinzip auch in möglichst ‚nahen‘ möglichen Welten in Kraft ist. Der naheliegendste Schluss ist laut Kim, dass sich die Gesamtsituation in etwa dergestalt rekonstruieren lässt: P verursacht P*, M superveniert auf P und M* superveniert auf P*. Die einzig genuin kausale Beziehung besteht also zwischen P und P*, die Beziehung zwischen M und M* hingegen muss als nicht-kausale Regularität begriffen werden.

Crisp und Warfield halten dieser Argumentation von Kim zwei Einwände entgegen. Der erste besteht in der Feststellung, dass Kim offensichtlich den Gedanken, dass jedes Vorkommen einer Geist-zu-Körper-Verursachung systematisch überdeterminiert ist, als sinnlos zurückweist.²²⁷ Crisp und Warfield, indessen, empfehlen denjenigen, die gleichzeitig an Eigenschaftsdualismus und Geschlossenheitsprinzip festhalten wollen, dass sie die Möglichkeit der

²²⁷ Vgl. Crisp, Thomas & Warfield, Ted: „*Kim's Master Argument*“, in: *Noûs*, Vol. 35:2, 2001, S. 312f.

systematischen Überdeterminierung sehr ernst nehmen sollten. Denn es ist trivial zu erkennen, dass die Konjunktion dieser beiden Prinzipien jeden Fall von Geist-zu-Körper-Verursachung kausal überdeterminiert. Der zweite Einwand von Crisp und Warfield richtet sich gegen Kims Argument, dass eine kausale Überdeterminierung und das Geschlossenheitsprinzip nicht miteinander vereinbar sind. Crisp und Warfield rekapitulieren und kritisieren Kims Argument wie folgt:

We're asked to consider a world W as much like ours as possible in which M occurs but P doesn't. Kim's suggestion here is that P* has a mental cause in W (namely, M) but no physical cause. The upshot: Closure is violated in W. But, thinks Kim, this isn't an acceptable consequence since surely a minimal counterfactual supposition like this wouldn't lead to such a major change in the world (...).

Consider though: either Supervenience holds in W or it does not. Suppose that it does. It follows that M has a physical supervenience base P' in W. What is the causal status of P' vis-à-vis P* in W? We won't repeat ourselves, but we saw above an argument of Kim's to the effect that if P' is a supervenience base of M and M causes P*, then P' is causally sufficient for P*. If Supervenience holds in W, therefore, P* does have a physical cause in W, and Closure therefore does not fail in W.

Now suppose that Supervenience does not hold in W. And suppose further that, just as Kim suggests, M causes P* in W without there being any physical cause of P*. Given these assumptions, Closure does indeed fail in W. But recall that we have supposed along with Kim that the actual world is a Supervenience world. It follows from this supposition that W is either nomologically or metaphysically impossible, depending on how we read the relevant modal operator in the formulation of Supervenience. So if W is a world in which Closure is violated in the way Kim suggests, W is at least nomologically impossible.²²⁸

Nach Crisp und Warfield gilt in der möglichen Welt W entweder Supervenienz oder sie gilt nicht. Wenn W also ein Welt ist, in der Supervenienz zutrifft, dann hat M – Kims Annahme entgegen – doch eine physikalische Supervenienzbasis, die allein schon kausal hinreichend für das Eintreten von P* ist. Die andere Möglichkeit ist, dass W eine Welt ist, in der die Supervenienz nicht zutrifft. Darunter würde jedoch sicherlich die Nähe von W zu der aktuellen Welt sehr leiden. Genauer gesagt würde es sich dann bei W sogar um eine nomologisch unmögliche Welt handeln – schließlich müsste in W Supervenienz gleichzeitig zutreffen und nicht zutreffen. Es ist zu erwarten, dass sich Anhänger der Idee der Überdeterminierung nun allerdings nicht von ihrer Vorstellung abbringen lassen, nur weil diese in einer (zumindest nomologisch) unmöglichen Welt nicht mit dem Geschlossenheitsprinzip vereinbar ist. Daraus ziehen Crisp und Warfield diese Schlussfolgerung: „We conclude, therefore, that Kim hasn't supplied us with a good reason for thinking that P and M can't be regarded as overdetermining causes of P*.“²²⁹

²²⁸ Crisp, Thomas & Warfield, Ted: „Kim's Master Argument“, in: *Noûs*, Vol. 35:2, 2001, S. 314.

²²⁹ Crisp, Thomas & Warfield, Ted: „Kim's Master Argument“, in: *Noûs*, Vol. 35:2, 2001, S. 314.

Versickern die kausalen Kräfte?

Ein weiterer Kritiker von Kims Exklusionsarguments ist Ned Block. In seinem 2003 in der Zeitschrift *Philosophy and Phenomenological Research* erschienenen Artikel *Do Causal Powers Drain Away?* argumentiert Block, dass, wenn Kims Exklusionsargument wahr wäre, kausale Kräfte sozusagen im Nichts versickern würden. Auch Block beginnt mit einer Rekonstruktion dessen, was das Exklusionsargument seiner Meinung nach ausdrückt. Wenn ein Ereignis ein anderes verursacht, sind dabei einige Eigenschaften kausal wirksam, andere wiederum nicht. Wenn beispielsweise eine Opernsängerin ein Glas mit dem Gesang des Wortes ‚Zerspringe!‘ zum bersten bringt, nimmt man wohl an, dass die Intensität und die Frequenz ihres Belcanto kausal wirksam waren und nicht etwa die Bedeutung des gesungenen Wortes. Genauso könnte für die Verursachung etwa auch die Dicke und Stabilität des Glases eine Rolle spielen, wohingegen der Ort, an dem das Glas gefertigt wurde, oder der Namen seines Besitzers unwichtig für die Kausalbeziehung wären. Wenn also eine Ursache aufgrund des Habens einer Eigenschaft F verursacht und die Wirkung aufgrund des Habens einer Eigenschaft F* verursacht wurde, stellt Block dies so da: $F \Rightarrow F^*$.²³⁰

Insgesamt hat das Exklusionsargument, Block zufolge, die Struktur einer *reductio ad absurdum*. Die erste Prämisse der *reductio* ist nun: $M \Rightarrow M^*$. Wenn nun gleichzeitig P die Supervenienzbasis von M und P* die Supervenienzbasis von M* ist, dann gilt auch: $P \Rightarrow P^*$. Daraus lässt sich folgendes Schema ableiten:

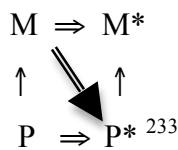
$$\begin{array}{ccc} M & \Rightarrow & M^* \\ \uparrow & & \uparrow \\ P & \Rightarrow & P^* \end{array} \quad ^{231}$$

Wobei die einfachen Pfeile für eine Supervenienzbeziehung und die doppelten Pfeile für eine Kausalbeziehung stehen. Das Zurückweisen der vermeintlichen Kausalbeziehung zwischen M und M* ist nun das Ziel von Kims *reductio*. Zunächst bemerkt Kim, dass es auf die Frage nach dem Eintreten von M* gemäß dieses Schemas zwei vollwertige Antworten gäbe und, dass zwischen beiden ein starkes Spannungsverhältnis bestehe. Denn M und P* scheinen eine

²³⁰ Vgl. Block, Ned: „*Do Causal Powers Drain Away?*“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 133f.

²³¹ Block, Ned: „*Do Causal Powers Drain Away?*“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 135.

jeweils vollständige und voneinander unabhängige Erklärung für das Eintreten von M^* zu liefern.²³² Daraus leitet Kim folgende Modifikation seines ursprünglichen Schemas ab:



M kann also M^* nur verursachen, indem es P^* verursacht. In diesem Fall konkurrieren tatsächlich zwei für sich alleine jeweils schon hinreichende Ursachen um die Verursachung von P^* . Daraus ergibt sich der Kern des Exklusionsarguments, den Block so beschreibt: „The crux of the Causal Exclusion Argument is Kim’s claim that P pre-empts M . To use a metaphor that Kim uses in a slightly different context, P does all the causal work, so there is nothing left over for M to do. Mental causation, Kim says, is threatened.“²³⁴ Aus dem besagten Spannungsverhältnis leitet Kim nun ab, dass P M bereits vorwegnimmt – durch die kausale Arbeit, die P verrichtet, wird M sozusagen „kausal arbeitslos“. Der zugrunde liegende Gedanke scheint zu sein, dass eine hinreichende Ursache *für sich allein* schon die ganze kausale Arbeit verrichtet. Da P nun zusätzlich auf einer grundlegenden Stufe als M steht, schließt die kausale Arbeit von P die von M aus. Allgemeiner gefasst resultiert daraus das Exklusionsprinzip, das Block so zusammenfasst: „(...) the principle that sufficient causation at one level excludes sufficient causation at another level (...).“²³⁵ Genau dieses Prinzip der kausalen Exklusion liegt im Fokus der Blockschen Kritik an Kim.

In Kims Argument spielt die Vorstellung der Überdeterminierung eine entscheidende Rolle: so nimmt Kim zunächst an, dass M und P zwei hinreichende Ursachen für P^* sind. Somit ist eine der aktuellen Welt möglichst nahe mögliche Welt W vorstellbar, in der M und P^* eintreten, wohingegen P in W nicht eintritt. In diesem Fall wäre M die alleinige Ursache von P^* . Dadurch würde aber wiederum das Prinzip der kausalen Abgeschlossenheit der Physik verletzt, denn es gäbe nun ein physikalisches Wirkungsereignis (P^*), das einzig und allein ein

²³² Ein reduktiver Eigenschaftsdualist, der Verursachung auf mehreren Ebenen annimmt, ist nach Block nicht gezwungen an dieser Stelle die Existenz eines Spannungsverhältnisses anzuerkennen – schließlich könnte er zugestehen, dass jedes Ereignis oberhalb der basalen Stufe der Physik horizontale Kausalbeziehungen und vertikale Determinierungsbeziehungen hat. Diese Erklärungen sind jedoch nicht völlig unabhängig voneinander. Um des Arguments willen nimmt Block aber mit Kim an, dass zwischen beiden Erklärungen ein Spannungsverhältnis besteht.

²³³ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 134.

²³⁴ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 135.

²³⁵ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 136.

nicht-physikalisches Ereignis (M) als Ursache hat. Die Hinfälligkeit des Geschlossenheitsprinzips ist nach Kim eine nicht hinnehmbare Folge. Aus diesem Grund muss die ursprüngliche Annahme ($M \Rightarrow M^*$) falsch sein – dies ist der letzte Schritt von Kims Argument, oder in anderen Worten das Vollziehen der *reductio*.

Grundsätzlich formuliert Ned Block zwei Einwände gegen den von Kim vertretenen Gedanken der Überdeterminierung: erstens setzt Kim in seinem Exklusionsargument voraus, dass in einer, der aktuellen Welt möglichst nahen, Welt W, in der P nicht eintritt, M trotzdem eintritt. Dies ist laut Block jedoch nicht unbedingt eine gesicherte Annahme. Entscheidend für das Gelingen oder Scheitern dieser Annahme ist das zugrundeliegende Verständnis von Supervenienz, weil, so Block, „(...) some supervenience doctrines but not others will preclude such a world. I don't see a way of choosing among these supervenience doctrines. (...) If there is no fact of the matter as to which of these supervenience doctrines holds, there is no fact of the matter of whether the closest non-P world contains an instance of M.“²³⁶ Nach Blocks Auffassung ist es also alles andere als klar, ob die der aktuellen Welt am nächsten gelegene mögliche Welt eine Welt sein kann, in der M eintritt und P sich nicht ereignet.

Das zweite Problem an Kims Vorstellung von Überdeterminierung besteht Block zufolge darin, dass man, selbst wenn man Kims Gedankenexperiment hinsichtlich der möglichen Welt folgt (also eine mögliche Welt, in der M eintritt und gleichzeitig P nicht eintritt, zugesteht), nicht zur der Annahme verpflichtet ist, dass diese mögliche Welt eine Welt ist, in der das Prinzip der kausalen Abgeschlossenheit verletzt wird. Dies illustriert Block an folgendem Beispiel:

Cain slew Abel by strangling him – let us suppose. This causal sequence was implemented by molecular trajectories, including the molecules of all or parts of Cain and Abel. Cain and Abel might have had a *merely* microphysical different diet for the previous ten years – that is, a diet of food that was not different from their actual food in a way that anyone could notice, but merely consisted of different tokens composed of different molecules – the same types of carbohydrates and proteins but different tokens of those types. Molecules in the food we eat and air we breathe are constantly replacing molecules in a human body. If Cain and Abel had eaten these different token molecules for ten years, they would have been composed mainly of different molecules, the strangling would have occurred just the same, but the molecules involved would have been different. Further, the strangling would have caused the death even if the molecular realizations had been different, so long as the difference wasn't significant at the macro level. So a world in which the molecular implementer of an event (token), e, does not exist can be a world in which e exists but is implemented differently. (...) But although there are worlds in which there is no alternative physical base, Kim owes us a reason for thinking the world in which there is no alternative physical base (in which causal closure fails) is closer than one in which it doesn't fail but M is implemented differently (or M doesn't exist at all).²³⁷

²³⁶ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 136.

²³⁷ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 136f.

Mit diesem Beispiel will Block zeigen, dass das Stattfinden eines gewissen Ereignisses e , das in der aktuellen Welt eine bestimmte Realisierung hat, in einer möglichen Welt W selbst dann noch kohärent denkbar ist, wenn es in W eine *andere* mikrophysikalische Implementierung hat. Hierzu nochmals kurz Blocks Beispiel: angenommen Kain hat Abel umgebracht, indem er ihn erwürgt hat. Nun ist eine mögliche Welt denkbar, in der Kain und Abel aus den gleichen Typen (aber auch aus anderen Token) physikalischer Grundbausteine bestehen. Obwohl diese Welt auf der mikrophysikalischen Ebene größtenteils von der aktuellen Welt divergieren würde, gäbe es keine Unterschiede auf der Makroebene – Kain würde Abel weiterhin erwürgen. Zudem hätte das Erwürgen auch den Tod Abels verursacht, selbst wenn es mikrophysikalisch *anders* (also nicht nur aus anderen Token, sondern auch aus anderen Typen von Molekülen) implementiert gewesen wäre – zumindest solange dies nicht zu gravierenden Unterschieden auf der Makroebene geführt hätte. Natürlich sind aber auch mögliche Welten vorstellbar, in denen es keine alternative physikalische Basis gibt. Ein Eintreten von P^* in einer solchen Welt würde tatsächlich das Geschlossenheitsprinzip verletzen, weil P^* dann nur noch eine nicht-physikalische Ursache – und zwar M – hätte. Laut Block bleibt Kim jedoch eine Antwort auf die Frage schuldig, warum eine solche Welt näher an der aktuellen Welt sein sollte als eine, in der das Geschlossenheitsprinzip erhalten bleibt und in der M anders realisiert ist (oder in der M gar nicht erst eintritt).

Ein weiteres Manko von Kims Exklusionsargument ist, dass es implizit eine Möglichkeit zur Identifikation von Ereignistoken über möglichen Welten hinweg voraussetzt. Vielleicht ist es einfach ungenau zu sagen, dass das Erwürgen auch stattgefunden hätte, wenn Täter und Opfer aus anderen Molekülen als in der aktuellen Welt zusammengesetzt gewesen wären. Schließlich ist es mehr als ungewiss, ob sich unter diesen Umständen genau *dieses bestimmte* Erwürgen ereignet hätte. Genau dies aber setzt Kim in seiner Formulierung des Exklusionsarguments voraus – wenn P^* in der möglichen Welt nicht stattgefunden hätte, M sich aber ereignet hätte, wäre auch M^* eingetreten – und zwar genau dasselbe M^* wie in der aktuellen Welt. Ein nicht-reduktiver Physikalist ist aber sicherlich nicht zu dieser Annahme verpflichtet. Vielmehr wäre es für ihn natürlicher anzunehmen, dass das Erwürgen nicht überhaupt nicht stattgefunden hätte, wenn Kain und Abel mikrophysikalisch anders realisiert gewesen wären, sondern dass dann eben ein *anderes* Erwürgen stattgefunden hätte. Es scheinen also gesetzesartige Zusammenhänge *gleichzeitig auf unterschiedlichen Ebenen* zu bestehen. Hierzu führt Block folgendes Beispiel ein: Zwischen Hurrikans und Zerstörung besteht eine gesetzesartige Beziehung. So hat etwa Hurrikan Edna große Zerstörung angerichtet bzw. verursacht. Nun hätten Ednas Luft- und Staubteilchen, sowie deren Flugbahn durchaus durch andere Teilchen (des gleichen

Typen) oder durch eine andere Flugbahn ersetzt werden können; trotzdem wäre das von Edna verursachte Ausmaß an Zerstörung gleich gewesen. Andererseits hätte es auch einen Unterschied auf der Mikroebene geben können, der einen Unterschied auf der Makroebene bedingt hätte – etwa wenn die Luftfeuchtigkeit höher gewesen wäre. Trotzdem hätte Edna auch in dieser alternativen mikrophysikalischen Realisierung einen ähnlichen Schaden anrichten können. Dies bringt Block zu dem Fazit: „If you think there are lawlike relations at many levels, and you think laws have something to do with the individuation of events and properties, you have some reason to endorse event individuation at levels other than the physical.“²³⁸ Wenn also gesetzesartige Zusammenhänge nicht nur auf der mikrophysikalischen Ebene anzusiedeln sind, ist zugleich anzunehmen, dass auch die Individuierung von Ereignissen auf anderen Ebenen als der mikrophysikalischen möglich sein muss.

Eine weitere Schwierigkeit von Kims Exklusionsproblem ist Block zufolge eng mit dem Exklusionsprinzip verquickt. Dieses birgt nämlich die akute Gefahr, dass kausale Kräfte sozusagen ins Bodenlose versickern. Hauptsächlich sieht Block das Exklusionsprinzip mit zwei Problemen konfrontiert: das erste Problem besteht darin, dass es, wenn das Exklusionsprinzip wahr wäre, keine mentale, keine physiologische, keine molekulare oder atomare Verursachung mehr gäbe. Schließlich wären Kausalrelationen *ausschließlich* auf den fundamentalsten Level der Physik beschränkt – die anderen Arten von Verursachung würden auf dieser Basisebene lediglich supervenieren und hätten daher selbst *keinerlei* kausale Kraft. Das zweite Problem hängt eng mit dem ersten zusammen. Denn für den Fall, dass das Exklusionsprinzip wahr ist und gleichzeitig keine Fundamentalebene der Physik existiert – Elementarteilchen also einfach unendlich teilbar sind – würden kausale Kräfte einfach ins Bodenlose versickern und es gäbe gleichwohl *überhaupt keine Verursachung* mehr. Laut Block ist aber die Frage nach einer untersten Ebene der Physik eine, nach aktuellem Stand der Wissenschaft, offene Frage: „There is some reason to think that the hypothesis that there is no bottom level of elementary particles is an open question from the point of view of the core of contemporary physical theory.“²³⁹ Natürlich ist diese Annahme nicht mit dem gängigen Standardmodell der Elementarteilchen, wonach es eine gewisse Anzahl verschiedener Elementarpartikel gibt, vereinbar. Dieses Modell scheint aber auch in der Wissenschaftsgemeinschaft umstritten zu

²³⁸ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 137f.

²³⁹ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 138.

sein.²⁴⁰ Folglich ist die Hypothese, dass es keine Fundamentalebene in der Physik gibt, dass Materie also unendlich teilbar ist und auf jeder dieser Ebenen unterschiedliche Eigenschaften zum Tragen kommen, eine offene Frage. ‚Offene Frage‘ bedeutet in diesem Zusammenhang weitaus mehr als eine bloße philosophische Möglichkeit oder eine mögliche skeptische Annahme – wie etwa die Möglichkeit, dass die gesamte Welt, mit all ihren ‚Beweisen‘ für ihre millionenjahrealte Geschichte, erst vor fünf Sekunden erschaffen wurde. Wenn es aber tatsächlich keinen Basislevel in der Physik gäbe und wenn gleichzeitig jede superveniente Eigenschaft ihre kausale Kraft ihrer jeweiligen subvenienten Eigenschaft verdanken würde, dann käme es zu einem Phänomen, dass man als *unendliche Subvenienz* bezeichnen könnte: Jede Behauptung der kausalen Wirksamkeit einer Eigenschaft würde durch die Existenz einer niedrigstufigeren Eigenschaft (auf die erstere superveniert) unterminiert werden – die kausalen Kräfte würden dann ins Bodenlose versickern und es gäbe überhaupt keine Verursachung in irgendeiner Form mehr.

Ein möglicher Einwand gegen diese Argumentation könnte sein, dass selbst wenn es eine offene Frage ist, ob es einen Fundamentallevel in der Physik gibt, es sicherlich keine offene Frage ist, ob es überhaupt Verursachung in irgendeiner Form gibt. Es mag zum Beispiel eine offene Frage sein, ob Rauchen Lungenkrebs verursacht, aber nicht ob irgendetwas etwas anderes verursacht. Aus diesem Grunde muss etwas an Kim Exklusionsargument fehlerhaft sein. Eine andere Art dies auszudrücken besteht nach Block im „(...) *Anti-Reductionist Conditional*: If there is no bottom level, then cancer never causes suffering or death.“²⁴¹ Ein gewiefter Anwalt, der die Interessen der Tabakindustrie oder eines Asbestherstellers vor Gericht vertritt, könnte nun mithilfe von Kims Exklusionsargument und diesem Prinzip argumentieren, dass seine Klienten nicht für Schadensersatzforderungen haftbar gemacht werden können. Schließlich hängt die Schuld der Zigaretten- und Asbestproduzenten dem anti-reduktionistischen Konditional zufolge von einer offenen Frage der Physik ab. Aus der offenen Frage nach der Existenz einer fundamentalen Ebene der Physik, wiederum, folgt die offene Frage, ob überhaupt irgendetwas etwas anderes verursacht und somit auch ob Rauchen bzw. Asbest (Lungen-)Krebs verursacht. Laut Block besteht der offensichtlichste Grund das Argument des Anwalts zurückzuweisen in folgendem Sachverhalt: „It is incompatible with our causal-explanatory practice (...) to regard the fact that it is an open question in physics whether there

²⁴⁰ Vgl. Block, Ned: „*Do Causal Powers Drain Away?*“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 138. Vor allem im zweiten Absatz nennt Block einige namhafte Wissenschaftler, die dem gängigen Standardmodell der Elementarteilchen widersprechen.

²⁴¹ Block, Ned: „*Do Causal Powers Drain Away?*“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 139.

is a bottom level as showing that it is an open question whether cancer ever causes death.“²⁴²

Es widerspreche schlichtweg der alltäglichen explanatorischen Praxis hinsichtlich Kausalrelationen, wenn man, wie der trickreiche Anwalt, von einer offenen Frage der Physik auf die Fragwürdigkeit von Ursache-Wirkung-Beziehungen im Allgemeinen schließen würde.

Nun könnte man an dieser Stelle auch einwenden, dass man wohl eher das anti-reduktionistische Konditional als Kims kausales Exklusionsargument zurückweisen müsse. Denn der Schritt von der Behauptung, dass Verursachung auf jeder Stufe durch eine niedrigere Verursachung unterminiert wird zu der Behauptung, dass es überhaupt keine Verursachung mehr gibt erscheint fragwürdig. Schließlich hängt die Gültigkeit dieses Gedankenganges entscheidend von einer offenen Frage hinsichtlich des Schlussfolgerns im Unendlichen und somit von einer offenen Frage der philosophischen Logik ab. Allerdings kann man das Argument leicht so modifizieren, dass es dieser Entgegnung standhält, so Block. Und zwar indem man das anti-reduktionistische Konditional durch folgendes Prinzip ersetzt: „(...) *Conservative Anti-Reductionist Conditional*: If there is no bottom level, and if the issue in philosophical logic concerning infinitary reasoning turns out as I suggested, then cancer never causes death.“²⁴³ Wenn nun Kims kausales Exklusionsargument wahr wäre, könnte der raffinierte Advokat der Tabakindustrie auch das konservative anti-reduktionistische Konditional verwenden, um seine Mandanten von jeglicher Schuld reinzuwaschen. Der Argumentationsgang des Anwalts wäre im Wesentlichen derselbe wie vorher, nur dass er dieses Mal zusätzlich annehmen würde, dass die Frage ob Krebs ursächlich für (die von der Anklage zitierten) Todes- oder Krankheitsfälle ist von *zwei* offenen Fragen abhängt – einer der Physik und einer der philosophischen Logik.

Von anderer Seite könnte eingewendet werden, die bloße Möglichkeit, dass es keine fundamentale Ebene in der Physik gibt, könne nicht gegen Kims Exklusionsargument verwendet werden. Sonst könnte analog in etwa so argumentiert werden: die bloße Möglichkeit, dass die Welt mit allen Indizien für ihre lange Geschichte erst vor fünf Sekunden erschaffen worden sein könnte, entzieht der vernünftigen Annahme, dass Krebs Todesfälle verursacht, ihren Boden. Denn zwischen dieser Annahme und dem konservativen anti-reduktionistischen Konditional bestehen keine essentiellen Unterschiede. Laut Block bestehen jedoch zwei erhebliche Unterschiede zwischen diesen beiden Hypothesen. Erstens mag es zwar durchaus eine metaphysische Möglichkeit sein, dass die Welt erst von fünf Sekunden erschaffen wurde – dies ist

²⁴² Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 139.

²⁴³ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 139.

allerdings keine offene Frage, sondern eine rein skeptische Behauptung. Zweitens würde die alltägliche Praxis via induktiven Schließens zur Annahme zu kommen, dass Krebs Todesfälle verursacht, durch die Behauptung die Welt sei erst vor fünf Sekunden komplett erschaffen worden, vollständig untergraben. Offensichtlicherweise könnte man sich nicht auf früheres Auftreten von Krebs- und Todesfällen berufen – es sei denn diese fanden in den letzten fünf Sekunden statt. Da es aber für induktive Schlüsse unerlässlich ist, sich auf (eine möglichst große Anzahl) frühere(r) Fälle zu beziehen, wäre diese Art der Schlussfolgerung in einer fünf Sekunden alten Welt einfach nicht mehr möglich. Die Annahme, dass es keine Fundamentalebene in der Physik gibt, hingegen, würde die Praxis des induktiven Schließens zunächst nicht unterminieren – es sei denn man nimmt gleichzeitig an, dass die kausalen Kräfte ausschließlich auf der untersten Ebene angesiedelt sind. Zudem gehörte die Annahme, dass Materie unendlich teilbar ist und es somit keine Fundamentalebene in der Physik gibt, lange Zeit zum Weltbild vieler gebildeter Menschen. Alles in allem gelangt Block schlussendlich zu folgendem Fazit: „In short, if there is no bottom level and if there is endless subvenience, then Kim’s Causal Exclusion Argument yields absurd results. The Exclusion Principle (that causally sufficient properties at one level exclude causally sufficient properties at another level) is to blame and should be rejected. Even if there is subvenience only down to elementary particle physics, there is a problem of causal powers draining down that far.“²⁴⁴

Gegen den Einwand des Versickerns der kausalen Kräfte wendet Kim laut Block folgende Verteidigungsstrategie an:

Kim’s defense against the line of thought I am advocating here (...) is that there is no problem of the causal powers of mental properties draining to the causal powers of the properties of cells, molecules, atoms or elementary particles, for these properties are all properties of the same thing, Socrates, and are therefore on the same level. He is using a notion of level keyed to objects. On this notion of level, the level of a property depends on what it is a property of. Properties of Socrates are on one level, properties of Socrates’ cells are on another level, properties of Socrates’ molecules are on another level, and so on.²⁴⁵

Nach Kims Verständnis ist die ‚Ebene‘ oder ‚Stufe‘ einer Eigenschaft also immer unmittelbar an das Objekt, das diese Eigenschaft exemplifiziert, geknüpft. Insofern werden die kausalen Kräfte mentaler Eigenschaften auch nicht an fundamentalere Ebenen (z.B. die zelluläre, molekulare oder elementare Ebene) abgegeben – denn schließlich sind dies alles Eigenschaften ein und desselben Dings (im Beispiel eben Sokrates). Es ist offensichtlich, dass nach diesem

²⁴⁴ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 140.

²⁴⁵ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 141f.

Verständnis von Ebene, die Ebene einer Eigenschaft im Wesentlichen von dem Objekt bzw. vom Träger dieser Eigenschaft abhängt. Allerdings existiert auch ein alternatives Verständnis von Ebene, so Block:

But there is another notion of level, which is keyed to *relations* among properties. The following is plausible: Socrates' pain supervenes on his neurological properties and his neurological properties supervene on the biochemical properties of his brain, and the biochemistry of his brain supervenes on the atomic-physical properties of his brain, and the atomic physics of his brain supervenes on the elementary-particle properties (if there are elementary particles) of his brain. The family of mental properties can be used to characterize the level of psychology, the family of neurological properties can be used to characterize the level of neuroscience, the family of elementary particle properties can be used to characterize the level of elementary particle physics. These distinct branches of science and their associated families of properties plausibly form a supervenience hierarchy (...). No mental difference without a neurological difference. No neurological difference without a bio-chemical difference. No bio-chemical difference without a difference in atomic physics. No difference in atomic physics without an elementary particle difference. And we can define a notion of level unlike Kim's that is keyed to branches of science. On Kim's notion of level, if the object of study is Socrates, the psychological, neurological, biochemical, atomic physics and elementary particle physics properties are all at the same level, despite belonging to different branches of science. On the notion of level I am suggesting (what Kim calls 'order'), Socrates' mental properties are at one level, his physiological properties are at another level and his atomic properties are still at another level.²⁴⁶

Das von Block verfochtene Verständnis von ‚Ebene‘ hängt hauptsächlich von den Beziehungen zwischen unterschiedlichen Eigenschaften ab. So kennzeichnen etwa mentale Eigenschaften die psychologische Ebene, neurologische Eigenschaften die Ebene der Neurowissenschaften, usw. Daraus ergibt sich eine Supervenienzhierarchie zwischen den unterschiedlichen Eigenschaften – so kann es beispielsweise keinen mentalen Unterschied ohne einen neurologischen Unterschied geben, keinen neurologischen Unterschied ohne einen biochemischen Unterschied, etc. Daraus lässt sich, Block zufolge, ein Verständnis von der Ebene einer Eigenschaft ableiten, das sich essentiell von Kims Auffassung unterscheidet. Dieser Interpretation von Ebene zufolge, ist die Ebene einer Eigenschaft nicht mehr an den Träger der Eigenschaft geknüpft, sondern ist vielmehr an einzelne Gebiete der Wissenschaft gebunden. Nach Kims Verständnis gehören alle Eigenschaften des Sokrates, also seine mentalen, neurologischen, biochemischen,... Eigenschaften *derselben* Ebene an, wohingegen sie nach Blocks Auffassung *unterschiedlichen* Ebenen zuzurechnen sind – die mentalen Eigenschaften Sokrates' gehören einer Ebene an, die neurologischen Eigenschaften einer anderen, etc. Gehören die Eigenschaften unterschiedlichen Ebenen an, besteht aber tatsächlich die Gefahr eines Versickerns der kausalen Kräfte auf die Fundamentalebene und wenn es keine solche Fundamentalebene gibt, dann drohen Kausalrelationen gänzlich zu verschwinden.

Nun könnte Kim argumentieren, dass es zunächst zwar den Anschein haben mag, mentale und physikalische Eigenschaften gehören unterschiedlichen Ebenen an, dies sich aber bei genaue-

²⁴⁶ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 142.

rem Betrachten als falsch herausstellt. Ein analoger Fall wäre etwa die Temperatur eines Gases und die Eigenschaften seiner Moleküle, beides sind unterschiedliche Eigenschaften eines Gases, gehören aber derselben Ebene an, so Kim. Für Block spielt dies jedoch eine eher untergeordnete Rolle, denn das Problem des Versickerns der kausalen Kräfte bleibt selbst dann bestehen, wenn man annimmt, dass die kausale Kraft der mentalen Eigenschaften des Sokrates zu den neurologischen Eigenschaften des Sokrates,... abwandert. Kim mag zwar gezeigt haben, dass die kausalen Kräfte nicht von großen Entitäten (Sokrates) zu kleineren Entitäten (Teile des Sokrates: z.B. seine Zellen, Moleküle,...) abfließen, aber das ist im Hinblick darauf, dass mentale Eigenschaften ihre kausale Kraft an die kausale Kraft der Eigenschaften von Elementarteilchen abgeben, nicht relevant. Im Wesentlichen geht es Block also um das Versickern der kausalen Kräfte von *Eigenschaften* zu tieferen Ebenen und nicht um das Versickern der kausalen Kräfte von größeren zu immer kleineren *Gegenständen*. Daher zieht Block folgende Bilanz: „I suppose that I should concede that at least Kim has plugged the causal drain from people to cells to molecules to elementary particles. But that is a Pyrrhic victory given that the draining issue concerning properties rather than things remains.“²⁴⁷

An dieser Stelle bringt Kim mikro-basierte Eigenschaften ins Spiel. Hierzu zunächst Blocks Definition dieser Eigenschaften: „(...) P is a micro-based property \equiv P is the property of being completely decomposable into non-overlapping parts $a_1 \dots a_n$ such that $P_1(a_1) \dots P_n(a_n) \ \& \ R(a_1 \dots a_n)$.“²⁴⁸ Was nach Kim das Versickern der kausalen Wirksamkeit stoppt, ist das Erreichen eben dieser mikro-basierten Eigenschaften in der Abwärtshierarchie. Die Besonderheit, die es den mikro-basierten Eigenschaften ermöglicht den Abfluss der kausalen Kräfte zu verhindern, ist die folgende: mikro-basierte Eigenschaften supervenieren auf bestimmten mereologischen Konfigurationen von Mikro-Eigenschaften. Aus diesem Grund besitzt die superveniente Eigenschaft auch kausale Kräfte, die kein einzelner Teil der subvenienten Konstellation (für sich) besitzt. Hierbei sichert die Identität zwischen mikro-basierten Eigenschaften und gewissen Mikro-Konfigurationen die kausale Wirksamkeit der mikro-basierten Eigenschaften ab. Falls z.B. ‚Wasser = H₂O‘ gilt, dann ist es völlig unproblematisch, wenn die kausale Wirksamkeit von Wasser, genauer gesagt der Eigenschaft Wasser-zu-sein, an H₂O abgegeben wird. Denn zu jedem Zeitpunkt, an dem H₂O kausal wirksam ist, ist automatisch auch Wasser kausal wirksam – dies beruht auf der Ununterscheidbarkeit von Identischem. Ähnlich verhält es sich auch mit mentalen Eigenschaften, deren kausale Rolle wird nämlich von mikro-

²⁴⁷ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 143.

²⁴⁸ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 143.

basierten Eigenschaften ausgefüllt. Diese mikro-basierten Eigenschaften sind nun identisch mit mereologischen Konstellationen von niedrigstufigeren Eigenschaften. Daraus folgt:

So the causal efficacy of the mental (outside of consciousness) does indeed drain down to the physiological, but the physiological doesn't drain down any further – at least not in any sense that makes the physiological inefficient – because of an identity between the physiological and the next “level” down, say biochemical. And the biochemical doesn't drain down any further because of an identity between the biochemical and say the atomic-physical. The mental is causally unreal, but the physiological is causally real and the draining stops there.²⁴⁹

Die kausalen Kräfte von mentalen Eigenschaften werden also tatsächlich ‚nach unten‘ an die physiologische Ebene abgegeben. Von dort werden sie allerdings nicht mehr weiter nach unten, sprich etwa an die biochemische oder atomar-physikalische Ebene, durchgereicht, zumindest nicht insoweit, als dass die physiologischen Eigenschaften dadurch ihre kausale Wirksamkeit verlieren würden. Schließlich sind die physiologischen Eigenschaften mit einer gewissen Menge und Zusammensetzung von biochemischen Eigenschaften identisch und bewahren genau durch diese Identität ihre kausale Wirksamkeit. Folglich ist das Mentale tatsächlich kausal nicht wirklich, das Physiologische hingegen schon – und an dieser Stelle stoppt auch der Abfluss der kausalen Kräfte.

Diesen Gedankengang hinterfragt Block folgendermaßen: „But why can't micro-based properties be micro-based in *alternative ways*? Why isn't jade an example of a micro-based property, micro-based in both calcium magnesium silicate (nephrite) and sodium aluminum silicate (jadeite)? (...) So my question is: why can't a micro-based property have more than one such decomposition?“²⁵⁰ Blocks zentrales Anliegen an dieser Stelle ist es aufzuzeigen, dass ein und dieselbe mikro-basierte Eigenschaft auch durchaus unterschiedlich realisiert sein kann. Im Fall von Jade etwa, tragen Mineralien mit jeweils unterschiedlichen Anteilen von Jadeit und Nephrit die Bezeichnung ‚Jade‘. So kann die mikro-basierte Eigenschaft ‚Jade-zu-sein‘ auf mehrere verschiedene Arten in jeweils unterschiedliche Elemente zerlegt werden. Wie bereits oben beschrieben, ist es ein kardinales Charakteristikum von mikro-basierten Eigenschaften, dass sie in verschiedene, nicht deckungsgleiche Bestandteile aufgespalten werden können. Daraus ergibt sich nach Block folgendes Problem: „Micro-based properties are supposed to prevent draining away for both supervenient and functional properties, but Kim's plugging the draining with micro-based properties depended on assuming identities (such as ‘water = H₂O’) and multiple composition will preclude such identities.“²⁵¹ Eine Grundvoraus-

²⁴⁹ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 145.

²⁵⁰ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 145.

²⁵¹ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 146.

setzung dafür, dass mikro-basierte Eigenschaften ein Versickern der kausalen Kräfte verhindern ist die Annahme von Identitäten (beispielsweise einer physiologischen Eigenschaft mit einer bestimmten Menge von biochemischen Eigenschaften, usw.). Sind mikro-basierte Eigenschaften nun aber auf unterschiedliche Art realisierbar, dann können auch keine Identitäten mehr postuliert werden. Folglich bricht das gesamte Konstrukt, das ein Abfließen der kausalen Kräfte verhindern sollte, in sich zusammen.

Um wiederum Blocks Argument der multiplen Zerlegbarkeit von mikro-basierten Eigenschaften anzugreifen, gibt es zwei grundsätzliche Möglichkeiten – zum einen kann der Aspekt des Multiplen unter Beschuss genommen werden, zum anderen kann man den Aspekt der Zerlegbarkeit ins Kreuzfeuer der Kritik nehmen. Kims Standardantwort auf die Behauptung eine höherstufige Eigenschaft U könne in zwei (und nur zwei) bestimmte, niedrigstufigere mereologische Konstellationen von Eigenschaften L_1 und L_2 multipel zerlegt werden, besteht in einer Fragmentierung der Makroebene. D.h. Kim würde nicht eingestehen, dass eine höherstufige Eigenschaft U auf zwei Arten zusammengesetzt sein kann, vielmehr würde er behaupten, es gebe in Wirklichkeit nicht eine, sondern vielmehr zwei höherstufige Eigenschaften, etwa U^* und U^{**} , von denen die eine mit L_1 und die andere mit L_2 identisch ist. Der Hintergrund dieser Idee Kims scheint zu sein, dass bei multipler Zerlegbarkeit unterschiedliche Gesetze am Werke sind und wo unterschiedliche Gesetze wirken, sind verschiedene natürliche Arten zu Gange. Diesen Gedanken testet Block an folgendem Beispiel: „The rigidity of a rigid body has a role in laws of motion, for example the conservation of angular momentum. This law explains for example why leaning while riding a bicycle causes it to turn. Rigid bodies have a number of decompositions, for example in a crystalline substance like diamond (a solid) and an amorphous substance like glass (a supercooled liquid). So rigidity appears to be both causally efficacious and multiply constituted – an affront to the fragmentation strategy.“²⁵² In diesem Beispiel scheint die Starrheit von Objekten einerseits kausal wirksam zu sein (z.B. im Drehimpulserhaltungssatz), andererseits ist die Starrheit unterschiedlicher Objekte (z.B. Diamant und Glas) aber auch in einer alternativen Zusammensetzung der Objekte begründet. Dies steht aber im offensichtlichen Widerspruch zur Fragmentierungsstrategie.

An späterer Stelle liefert Block eine Rekonstruktion von Kims Problem aus seiner Sichtweise:

The problem for Kim could be put in the following way: Suppose we have good reason to believe all of the following:

1. $U_a \Rightarrow U_b$ is a causal law
2. U_a has alternative decompositions in mereological configurations L_1 and L_2

²⁵² Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 147.

3. L_1 and L_2 participate in different laws.

U_a would appear to be both multiply constituted and causally efficacious in violation of the fragmentation strategy. If L_1 and L_2 participate in different laws, that counts against regarding U_a as simply identical to the disjunction of L_1 and L_2 , for the disjunction of L_1 and L_2 would not be a genuine property (from Kim's perspective). But the option of regarding U_a as fragmenting into U_a^* and U_a^{**} would seem unpalatable (though perhaps not to Kim) because of the fact that U_a itself participates in a causal law. Kim seems to regard the nomological claims of L_1 and L_2 as undercutting the nomological claims of U_a – that seems the upshot of the point of view behind the fragmentation strategy. The idea would be that the separate laws of L_1 and L_2 show that there are no real laws of U_a but only similar laws of L_1 and L_2 . That is, rather than the law ' $U_a \Rightarrow U_b$ ', Kim would suppose (on this interpretation) that there are two laws, ' $L_1 \Rightarrow U_b$ ' and ' $L_2 \Rightarrow U_b$ '. If this is Kim's view, he owes us a rationale for it.²⁵³

Kims Fragmentierungsstrategie beruht also im Wesentlichen auf der Behauptung, es handle sich nicht um ein, sondern um zwei unterschiedliche höherstufige Eigenschaften, bei denen jeweils ein eigenes, wenngleich auch sehr ähnliches, Gesetz am Werke ist. Warum aber zwei unterschiedliche Gesetze wirken sollten, dafür liefert Kim, Block zufolge, jedoch keine vernünftige Begründung. Sollte die Begründung nämlich die sein, dass auf diese Weise einige Probleme mentaler Verursachung gelöst werden können, wäre das nicht zwangsläufig eine überzeugende Argumentation. Schließlich hätte ein konkurrierendes Prinzip den gleichen Anspruch auf Gültigkeit – und zwar: Versteht man kausale Wirksamkeit als ein Phänomen, das gleichzeitig auf mehreren Ebenen präsent sein kann und betrachtet die unterschiedlichen Ebenen nicht als konkurrierend, würden damit auch Probleme mentaler Verursachung gelöst. Zudem hätte letztere Begründung den Vorteil, dass damit auch das Problem des Versickerns kausaler Kräfte vermieden werden könnte.

Alles in Allem zieht Block folgendes, abschließendes Fazit hinsichtlich Kims Exklusionsproblem: „In short, Kim uses the fragmentation strategy to shore up the Exclusion Principle, but the fragmentation strategy remains to be justified. The Exclusion Principle (and the Causal Exclusion Argument that depends on it) appears to yield conclusions that conflict with our explanatory practice. This conflict entitles us – and causes us – to reject the Causal Exclusion Argument until these justifications are provided.“²⁵⁴

Überdeterminierung II

Auch von Seiten Ausonio Marras' erfährt Kims Exklusionsargument harsche Kritik. In seinem im Jahr 2007 in der Zeitschrift *Erkenntnis* publizierten Artikel *Kim's Supervenience Ar-*

²⁵³ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 149.

²⁵⁴ Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003, S. 149f.

gument and Nonreductive Physicalism greift Marras das Exklusionsargument an. Für ihn ist das Exklusionsargument kein Argument gegen den Physikalismus generell, sondern ist gegen eine spezielle Spielart des Physikalismus – nämlich den nicht-reduktiven Physikalismus – gerichtet. Er versucht mitunter zu zeigen, dass das Exklusionsargument, dessen zweite Stufe die Struktur einer *reductio ad absurdum* aufweist, deshalb nicht erfolgreich ist, weil für die *reductio* entscheidende Voraussetzungen nicht gegeben sind.²⁵⁵ Hierzu stellt Marras Kims Argumentationsgang nochmals kurz in eigenen Worten dar:

The second stage of Kim's argument, recall, turned on the question of what caused P^* : was it M or was it P ? This part of the argument takes the form of a *reductio*: (1) Suppose the cause of P^* was M ; but then, (2) by [Closure], there must have also been a co-occurrent physical cause of P^* , which we may suppose was P ; and (3) assuming that this was not a genuine case of overdetermination, we must, by [Exclusion], deny that both M and P caused P^* , unless of course M and P were not synchronous and distinct; (4) but M and P were synchronous and, given [Irreducibility], they were distinct; (5) hence P^* can't have had both M and P as its cause; (6) hence, on pain of giving up [Closure], the real cause of P^* must have been P and not M . *QED*.²⁵⁶

Die epiphanomenalistische Schlussfolgerung (6) ist der Preis, den ein nicht-reduktiver Physikalist für seine Vorliebe für das Irreduzibilitätsprinzip bezahlen muss, so Kim. Da nämlich M und P^* voneinander verschieden und somit auch nicht aufeinander reduzierbar sind, muss eines von beiden aus der Kausalerklärung für P herausfallen. Und dem Geschlossenheitsprinzip zufolge kann die Entscheidung, ob M oder P^* eliminiert werden muss, nur auf M fallen. Es ist offensichtlich, dass das Exklusionsprinzip eine gewichtige Rolle in Kims Argument spielt. Marras gibt das Exklusionsprinzip als quantifiziertes Konditional wieder:

[Ex] For all events c , c' , and e : if (i) c is synchronous with c' ; (ii) $c \neq c'$; and (iii) e is not ("genuinely") overdetermined by c and c' , then it is not the case that both c is a sufficient cause of e and that c' is a sufficient cause of e .²⁵⁷

Anschließend stellt Marras M , P und P^* auf den Prüfstein, ob sie den oben genannten Bedingungen des Exklusionsprinzips genügen. Dabei hegt Marras jedoch starke Zweifel, ob sie alle Kriterien erfüllen – insbesondere fokussiert Marras seine Kritik hierbei auf die Konditionen (ii) und (iii). De facto wäre es auch wenig sinnvoll die Klausel (i) vor diesem Hintergrund anzugreifen, weil sich M und P nämlich per Supervenienzbedingung gleichzeitig ereignen müssen.

²⁵⁵ In seinem Essay nennt Marras zwei weitere Kritikpunkte an Kims Exklusionsargument, die an dieser Stelle jedoch nicht vorgestellt werden.

²⁵⁶ Marras, Ausonio: „Kim's Supervenience Argument and Nonreductive Physicalism“ in: *Erkenntnis*, Vol. 66, 2007, S. 315.

²⁵⁷ Marras, Ausonio: „Kim's Supervenience Argument and Nonreductive Physicalism“ in: *Erkenntnis*, Vol. 66, 2007, S. 316.

Zunächst also zu Bedingung (ii): Angenommen M und P kommen beide als Ursache von P* in Frage. Nun ist es zunächst ungewiss, ob sie auch wirklich verschieden voneinander sind. Dies scheint zwar zunächst über jeden Widerspruch erhaben zu sein – schließlich dürfte ein nicht-reduktiver Physikalist wohl kaum geneigt sein das Irreduzibilitätsprinzip aufzugeben. Allerdings sollte man sich in diesem Zusammenhang Folgendes vor Augen führen: in Kims Exklusionsargument figurieren M, P und P* als Ursachen bzw. Wirkungen. Zwar ist von ihnen (in verkürzter Redeweise) als Eigenschaften die Rede, aber es handelt sich bei M, P und P* jedoch nicht um *Eigenschaften*, sondern vielmehr um *Instanziierungen von Eigenschaften*. Allerdings haben Instanziierungen von Eigenschaften in Kausalrelationen sicherlich den Status von Partikularen. Ein Grundsatz von vielen nicht-reduktiven, physikalistischen Ansätzen ist aber, dass während mentale Eigenschaften nicht mit physikalischen Eigenschaften identisch sind, mentale Partikulare mit physikalischen Einzeldingen bzw. -ereignissen sehr wohl identisch sein können – dies ist geradezu der Kern des Token-Physikalismus. Die Verschiedenheit von mentalen und physikalischen *Eigenschaften*, die im Irreduzibilitätsprinzip ihren Ausdruck findet, hat jedoch nicht die Verschiedenheit von mentalen und physikalischen *Partikularen* zur Folge. Eben dies wird aber im Schritt (4) der *reductio* vorausgesetzt. Dementsprechend muss ein nicht-reduktiver Physikalist diesen Schritt auch nicht mitgehen, denn „(...) it assumes that [Irreducibility] entails the distinctness of M and P *qua property instance*. But then such property instances must be rather different sorts of entities from kinds of particulars these nonreductive physicalists typically have in mind as terms of causal relations. What kind of entities are, then, Kim’s property instances?“²⁵⁸ Marras zufolge entsprechen Instanziierungen von Eigenschaften an dieser Stelle am ehesten dem, was Kim unter einem ‚Ereignis‘ versteht – also einer komplexen, zusammengesetzten Entität mit der Struktur $[x, P, t]$, wobei die Elemente der Struktur ein Individuum x, eine konstituierende Eigenschaft P und ein Zeitpunkt t sind. Im Beispielfall exemplifiziert ein Individuum x zum Zeitpunkt t die Eigenschaft P. Aus diesem Verständnis von Eigenschaftsinstanziierungen lässt sich folgende Identitätsbedingung für zwei verschiedene Ereignisse ableiten: $[x, F, t] = [x, G, t] \rightarrow F = G$. Zwei Instanziierungen von Eigenschaften sind also genau dann identisch, wenn die beiden Eigenschaften identisch sind (und zum selbem Zeitpunkt im gleichen Individuum instanziiert werden). Mit dieser Identitätsbedingung fallen jedoch Token- und Type-Identität zusammen. Diese Unterscheidung ist aber für viele Vertreter von nicht-reduktiven, physikalistischen Ansätzen zentral – gerade durch diesen Aspekt grenzen sich viele nicht-reduktive Physikalisten

²⁵⁸ Marras, Ausonio: „Kim’s Supervenience Argument and Nonreductive Physicalism“ in: *Erkenntnis*, Vol. 66, 2007, S. 316.

nach ihrem Verständnis von reduktiven Physikalisten ab. Aus diesem Grund kann auch Schritt (4) der *reductio* als missglückt zurückgewiesen werden. Schließlich müssen M und P nicht zwangsläufig die Bedingung (ii) des Exklusionsprinzips erfüllen. Denn M und P fungieren qua Instanziierung von Eigenschaften als Partikulare in der Kausalbeziehung mit P*. Daher kann auch ein nicht-reduktiver Physikalist durchaus behaupten, M und P seien miteinander identisch (natürlich würde er aber weiterhin bestreiten, dass die Eigenschaft-M-zu-sein mit der Eigenschaft-P-zu-sein identisch ist). Damit büßt auch das Exklusionsargument maßgeblich an Überzeugungskraft ein und kann somit von nicht-reduktiven Physikalisten zurückgewiesen werden. Dieser rapide Verlust an Glaubwürdigkeit ist allem voran darauf zurückzuführen, dass bereits das Verständnis wie Eigenschaften gemäß Kims Exklusionsargument instanziiert werden, eine implizite Ablehnung des nicht-reduktiven Physikalismus beinhaltet. Daher muss das Exklusionsargument im strengen Sinne weder auf das Exklusions- noch auf das Geschlossenheitsprinzip zurückgreifen, um den Token-Physikalismus zu widerlegen. Nun ist es aber offensichtlich, dass ein Argument, das im Wesentlichen gegen eine metaphysische Position gerichtet ist, nicht deren Negation bereits als gegeben voraussetzen darf. Oder wie Marras es formuliert: „To someone who regards some such version of nonreductive physicalism as (part of) a substantive thesis and not as a thesis refutable simply by definition, Kim’s supervenience argument is thus bound to appear profoundly question begging.“²⁵⁹ Andererseits gesteht Marras aber auch zu, dass es andere Versionen des nicht-reduktiven Physikalismus gibt, die mit Kims feinkörniger Konzeption von Ereignissen und deren Identitätsbedingungen einhergehen. Um dem zentralen Gedanken der Irreduzibilität dennoch gerecht zu werden, postulieren solche Theorien, dass mentale Ereignisse von physikalischen Ereignissen realisiert werden bzw. darin begründet sind (anstatt token-identisch mit ihnen zu sein). Gibt man die Identität zwischen mentalen und physikalischen Ereignistoken jedoch auf, stellt sich die Frage, ob ein Ansatz, demzufolge nicht *alle* Partikulare einer Kausalbeziehung *rein* physikalisch sind, noch als Physikalismus bezeichnet werden kann. Für diese streitbare These äußert Kim jedoch kein Argument, obwohl er mit dem Exklusionsargument stillschweigend vorauszusetzen scheint, dass gerade solche nicht-reduktive, physikalistische Theorien ein angemessenes und gut gewähltes Ziel für seine Kritik sind.

Aber selbst für den Fall, dass man Kim zugesteht Bedingung (ii) des Exklusionsprinzips werde von M, P und P* erfüllt, bleibt weiterhin die Frage offen, wie es um Bedingung (iii) steht. Um dem Exklusionsprinzip gerecht zu werden, dürfen (die Instanziierungen von) M und P demnach (die Instanziierung von) P* nicht genuin überdeterminieren. Falls es sich nämlich

²⁵⁹ Marras, Ausonio: „Kim’s Supervenience Argument and Nonreductive Physicalism“ in: *Erkenntnis*, Vol. 66, 2007, S. 317.

um einen solchen Fall genuiner Überdeterminierung handeln würde, wäre dadurch das Exklusionsprinzip nicht verletzt (und infolgedessen wäre die gesamte *reductio* nicht erfolgreich). In diesem Zusammenhang ist es zunächst interessant nachzufragen, was Kim eigentlich unter genuiner Überdeterminierung versteht. Ein ‚gewöhnlicher‘ Fall von Überdeterminierung ist etwa das altbekannte Zwei-Kugeln-Beispiel: Eine Person wird zeitgleich von zwei Projektilen getroffen, wobei jede Kugel für sich allein genommen schon hinreichend für den Tod der Person gewesen wäre. Das Hauptmerkmal solcher ‚gewöhnlicher‘ Fälle von Überdeterminierung ist laut Kim, dass sich zwei voneinander unabhängige Kausalketten schneiden und in einer gemeinsamen Wirkung münden. Marras gesteht Kim zu, dass er sicherlich richtig damit liegt, M und P nicht als überdeterminierend in diesem Sinne aufzufassen. Eine solche Omnipräsenz von fortwährender Überdeterminierung in der natürlichen Ordnung wäre jedem Naturalisten sicherlich zutiefst zuwider und müsste als rätselhaft und bizarr erachtet werden. Es ist einfach zu sehen, dass dieser Form von gewöhnlicher Überdeterminierung eine Bedingung kontrafaktischer Unabhängigkeit immanent ist. Das überdeterminierte Ereignis hätte nämlich selbst dann stattgefunden, wenn eine der beiden Ursachen nicht eingetreten wäre. Im Bezug auf M und P würde das bedeuten, dass sie P* nur dann unabhängig überdeterminieren würden, wenn gelten würde (a) wenn M eingetreten wäre und P nicht, wäre P* trotzdem eingetreten und (b) wenn P eingetreten wäre und M nicht, wäre P* trotzdem eingetreten. Da zumindest eine dieser Bedingungen falsch sein muss, kann es sich (nach diesem Verständnis) nicht um einen Fall genuiner Überdeterminierung handeln. Hiergegen wendet Marras folgenden Punkt ein:

However, one might wonder, why should non-ordinary, *dependent* overdetermination be any less “genuine” than *independent* overdetermination, as long as it is non-mysterious – resulting not from chance or design, but from the natural order (as it would in worlds in which supervenience obtains)? Or, to put it less tendentiously, why should it matter whether the two allegedly sufficient causes are counterfactually independent of each other or not? If an event can have two sufficient causes one of which depends on the other, then an event can have two sufficient causes. So it is hard to see why Kim should exempt only independent overdetermining causes from exclusion, and not dependent ones as well.²⁶⁰

In der Tat ist es fraglich, warum Fälle genuiner Überdeterminierung sich auf diejenigen Begebenheiten beschränken sollten, bei denen die hinreichenden Ursachen kontrafaktisch voneinander unabhängig sind. Was ist daran mehr ‚genuin‘ als an Umständen, bei denen die beiden überdeterminierenden Ursachen voneinander abhängig sind? Damit ist nicht ersichtlich, weshalb solche Fälle nicht – genauso wie die Vorkommnisse, bei denen die beiden hinreichenden Ursachen kontrafaktisch voneinander unabhängig sind – auch vom Exklusionsprinzip ausgenommen werden sollten. Allerdings kann Supervenienz als Paradebeispiel für eine Form de-

²⁶⁰ Marras, Ausonio: „Kim’s Supervenience Argument and Nonreductive Physicalism“ in: *Erkenntnis*, Vol. 66, 2007, S. 318f.

pendenter Überdeterminierung erachtet werden. Folglich ist es unklar warum Fälle, bei denen eine hinreichende Ursache auf einer anderen hinreichenden Ursache superveniert, nicht auch vom Exklusionsprinzip ausgenommen werden sollten. Zwar geht Marras mit Kim davon aus, dass zwei überdeterminierende Ursachen generell voneinander unabhängig sein sollten, dies spiele aber für das Klären der Frage, ob das Exklusionsprinzip verletzt wird, wenn M und P beide als vermeintliche Ursachen von P* gehandelt werden, eigentlich keine Rolle. Schließlich stünde man ohnehin vor folgendem Dilemma:

For, suppose that the independence condition *is not* part of the concept of overdetermination: then, as just seen, there seems to be no good reason to deny that *M* and *P* can be overdetermining causes of *P**, in which case we have no violation of the [Exclusion] principle. Suppose instead that the independence condition *is* built into the very concept of overdetermination; then *M* and *P* are not overdetermining causes, as they do not cause *P** *independently* of each other. But for the same reason, it can be argued, it is *not true* that each qualifies as a *sufficient* cause of *P**, as [Exclusion] requires, so, again, there is no violation of [Exclusion]. The reason for this is not difficult to see. Ordinarily, when we say that a cause *c* is sufficient for an effect *e* (under the prevailing circumstances), we imply that *no other condition is necessary* – whence the intuitive plausibility of [Exclusion].²⁶¹

Für das Exklusionsargument scheint es also nicht relevant zu sein, ob man die Unabhängigkeit zweier überdeterminierender Ursachen als inhärent für das Konzept von Überdeterminierung betrachtet. Wenn die Unabhängigkeitsbehauptung nämlich nicht ein integraler Bestandteil der Vorstellung von Überdeterminierung ist, dann könnten zwei hinreichende (überdeterminierende) Ursachen auch voneinander abhängig sein, z.B. indem die eine auf die andere superveniert. Im Bezug auf M und P würde dies bedeuten, dass beide P* überdeterminieren könnten, ohne dass dabei M oder P vom Ausschluss aus der Kausalrelation mit P* bedroht wären. Schließlich würde im Falle genuiner Überdeterminierung – und ein solcher wäre bei diesem Verständnis von genuiner Überdeterminierung gegeben – die kausale Wirksamkeit keiner der beiden hinreichenden Ursachen durch das Exklusionsprinzip in Frage gestellt. Wenn andererseits die These vertreten wird, dass zwei hinreichende Ursachen voneinander unabhängig sein müssen, um ein Ereignis zu überdeterminieren, dann ist es offensichtlich, dass M und P in diesem Fall P* nicht überdeterminieren würden, weil sie P* nicht unabhängig voneinander verursachen. Aus dem selben Grund könnte man aber auch argumentieren, dass M und P deswegen nicht als zwei jeweils für sich hinreichende Ursachen von P* gelten können. Schließlich würde das alleinige Eintreten von M oder P nicht zum Eintreten von P* führen. Sind aber M und P nicht zwei hinreichende Ursachen von P*, so wäre auch ihre jeweilige kausale Wirksamkeit nicht vom Exklusionsprinzip bedroht. Marras vertritt die Auffassung,

²⁶¹ Marras, Ausonio: „*Kim's Supervenience Argument and Nonreductive Physicalism*“ in: *Erkenntnis*, Vol. 66, 2007, S. 319.

dass M und P qua Supervenienzbedingung voneinander abhängig sind. Daher gelangt er auch zum abschließenden Fazit:

Given [Supervenience], it is obviously not possible for *M* to occur without its supervenience base (*P*, on this occasion) occurring, or for its supervenience base to occur without *M* occurring. In any nomologically possible world, the occurrence of a supervenient property on a given occasion necessitates, and is necessitated by, the occurrence of (whatever may be) its base on that occasion. But then, if neither *M* nor *P* can *occur* in absence of the other, neither can *cause P** in the absence of the other: *P* is causally sufficient for *P** if and only if *M* is. Thus neither *M* nor *P* qualifies as an independently sufficient cause of *P** in the required exclusionary sense. So [Exclusion] is not violated, and the supervenience argument fails.²⁶²

Aufgrund der Supervenienzbedingung gilt Marras zufolge in jeder nomologisch möglichen Welt: Genau dann, wenn P instanziiert wird, wird auch M instanziiert und umgekehrt. Folglich kann in keiner nomologisch möglichen Welt P ohne M bzw. M ohne P auftreten. Wenn aber beide Ereignisse nicht unabhängig voneinander eintreten können, dann können sie gewiss auch nicht unabhängig vom jeweils anderen P* verursachen. Eben dies ist aber eine grundlegende Prämisse des Exklusionsprinzips. Da das Exklusionsprinzip an dieser Stelle versagt, stürzt somit das gesamte Exklusionsargument in sich zusammen.

Abschließend präsentiert Marras einen weiteren Punkt, den er am Exklusionsargument bemängelt. In erster Linie richtet sich das Exklusionsargument, wie gesagt, gegen nicht-reduktive Physikalisten, allerdings trifft es nur einen Teil der intendierten Zielgruppe, denn „(...) it would, in any case, fail to move those nonreductive physicalists who do not see epiphenomenalism necessarily as a threat, but as something which, like it or not, we need to incorporate within our most defensible (if not most cherished) world picture.“²⁶³ Kim ist sich dessen auch bewusst und möchte daher den Epiphänomenalismus nicht von vornherein ausschließen. Vielmehr möchte er nicht-reduktive Physikalisten mit dem Exklusionsargument zwingen entweder den nicht-reduktiven Physikalismus fallen zu lassen oder aber eine epiphänomenale Version zu vertreten. Zwar scheint Kim epiphänomenalistisches Gedankengut als sehr beunruhigend zu empfinden, gibt aber keinen expliziten Grund dafür an, warum man es geringschätzen sollte. Zudem vertritt Kim im Hinblick auf phänomenale Eigenschaften, also auf Qualia, an anderer Stelle genau die Ansicht, dass jene superveniente und nicht reduzierbare Eigenschaften sind, zugleich aber keinerlei kausale Kraft besitzen oder in anderen Worten epiphänomenale Eigenschaften sind. Darum ist es umso weniger klar, warum Schmerzen etwa epiphänomenal sein sollten, während Intentionen es nicht sind. Dazu Marras:

²⁶² Marras, Ausonio: „*Kim's Supervenience Argument and Nonreductive Physicalism*“ in: *Erkenntnis*, Vol. 66, 2007, S. 319.

²⁶³ Marras, Ausonio: „*Kim's Supervenience Argument and Nonreductive Physicalism*“ in: *Erkenntnis*, Vol. 66, 2007, S. 320.

It would indeed be ironic if Kim were to deny the nonreductive physicalist the general strategy that he himself adopts with respect to phenomenal properties: i.e. *first* determine whether a given class of properties is physically reducible or irreducible, *then* determine what may follow from this about their causal powers (...). A nonreductive physicalist who had come to a prior determination that all mental properties are indeed irreducible could well make use of Kim's supervenience argument and be entitled to conclude (were the argument sound) that *all* mental properties – belief no less than pain – are epiphenomenal. In other words, Kim can't have it both ways: he cannot both acquiesce in his position which allows phenomenal properties to be causally impotent (if real at all), and at the same time chide a nonreductive physicalist who accepted the supervenience argument for condoning the causal impotence of mental properties in general.²⁶⁴

Marras wirft Kim also vor sozusagen mit zweierlei Maß zu messen: Denn während er einerseits selbst argumentiert phänomenale mentale Eigenschaften seien nicht reduzierbar auf physikalische Eigenschaften, zugleich aber auch kausal wirkungslos, scheint er andererseits nicht-reduktiven Physikalisten die Sichtweise, dass alle mentalen Eigenschaften epiphänomenal sind, nicht zuzugestehen. Laut Marras stünde es nicht-reduktiven Physikalisten aber jedoch durchaus zu, das Exklusionsargument (wenn es stimmig wäre) zu verwenden, um zu genau diesem Schluss (dass alle mentalen Eigenschaften epiphänomenal sind) zu gelangen.

²⁶⁴ Marras, Ausonio: „*Kim's Supervenience Argument and Nonreductive Physicalism*“ in: *Erkenntnis*, Vol. 66, 2007, S. 321.

Mögliche Lösungsstrategien für das Exklusionsproblem

Determinablen und Determinaten

An den Anfang seines Lösungsvorschlages für das Exklusionsproblem stellt Yablo zunächst eine kurze Darstellung des Exklusionsarguments aus seiner Sicht:

“How can mental phenomena affect what happens physically? Every physical outcome is causally assured already by preexisting physical circumstances; its mental antecedens are therefore left with nothing further to contribute.” This is the *exclusion argument* for epiphenomenalism. Here is the argument as it applies to mental events; for the version which applies to properties, replace ‘event x ’ with ‘property X ’:

- (1) If an event x is causally sufficient for an event y , then no event x^* distinct from x is causally relevant to y (*exclusion*).
- (2) For every physical event y , some physical event x is causally sufficient for y (*physical determinism*).
- (3) For every physical event x and mental event x^* , x is distinct from x^* (*dualism*).
- (4) So: for every physical event y , no mental event x^* is causally relevant to y (*epiphenomenalism*).

This is bad enough (...) but a simple extension of the argument promises to deprive mental phenomena of all causal influence whatsoever. Every event z of whatever type is metaphysically necessitated by some underlying physical event y , whose causally sufficient physical antecedents are presumably sufficient for z as well. But then by the exclusion principle, z ’s mental antecedents are irrelevant for its occurrence. So, mental phenomena are *absolutely* causally inert. And now it is not only speech and action that are chimerical but also thinking.

Note well that the exclusion argument raises *two* problems for mental causation, one about mental particulars (events), the other about mental properties. Strangely, philosophers have tended to treat these problems in isolation and to favor different strategies of solution.²⁶⁵

Für Yablo ist das Exklusionsargument also in erster Linie ein Argument, das unter bestimmten metaphysischen Voraussetzungen, zu einer epiphänomenalistischen Sichtweise zwingt. Über die Akzeptanz von drei Prämissen – die Yablo ‚Exklusion‘, ‚physikalischer Determinismus‘ und ‚Dualismus‘ nennt – gelangt es zu seiner Konklusion, dass das Mentale nur einen epiphänomenalen Charakter besitzt. Wobei sich das Argument auch sehr einfach so ausweiten lässt, dass sich mentale Phänomene jeglicher Art, d.h. nicht nur Sprechen und Handlungen, sondern etwa auch das Denken als kausal wirkungslos erweisen, so Yablo. Denn jedes Ereignis z im Universum besitzt ein physikalisches Antezedens, das für sich allein genommen notwendig und (vermutlich) auch hinreichend für das Eintreten von z ist. Dementsprechend verliert das Mentale jegliche kausale Wirksamkeit. Hinsichtlich mentaler Verursachung eröffnet das Exklusionsargument laut Yablo genau genommen sogar zwei Probleme – und zwar eines bezüglich der kausalen Wirksamkeit mentaler Partikulare (i.e. Ereignisse) und ein zweites hinblicklich der kausalen Relevanz mentaler Eigenschaften.

²⁶⁵ Yablo, Stephen: „*Mental Causation*“ in: *The Philosophical Review*, Vol. 101, 1992, S. 246ff.

Der Lösungsweg, der im Anbetracht des ersten Problems – also der kausalen Wirksamkeit mentaler Ereignisse – üblicherweise beschritten wird, besagt, dass mentale Ereignisse kausale Kraft besitzen, weil sie mit physikalischen Ereignissen identisch sind. Dies verschiebt das Problem der kausalen Relevanz jedoch auf die Ebene der Eigenschaften. Denn sollte ein mentales Ereignis aufgrund seiner Identität mit einem physikalischen Ereignis kausal wirksam sein, kann es nicht gleichzeitig aufgrund seines Besitzes einer bestimmten mentalen Eigenschaft kausal wirksam sein, sondern muss seine kausale Relevanz allein seinen physikalischen Eigenschaften verdanken. Aufgrund der multiplen Realisierbarkeit des Mentalen erachtet Yablo eine Identifizierung von mentalen Eigenschaften mit physikalischen Eigenschaften als unmöglich. Denn um identisch zu sein, müssen sich Eigenschaften gegenseitig notwendigerweise bedingen. Allerdings ist jede physikalische Eigenschaft, die spezifisch genug ist eine mentale Eigenschaft notwendigerweise zu bedingen, so spezifisch, dass eine notwendige Bedingung in der Gegenrichtung scheitern muss.²⁶⁶ So ist etwa der neuronale Zustand N, in dem sich eine Person, beispielsweise Hans, befindet, wenn er einen stechenden Schmerz im kleinen Finger seiner rechten Hand empfindet keine notwendige Bedingung für das Empfinden stechender Schmerzen im kleinen Finger der rechten Hand. Es ist beispielsweise durchaus denkbar, dass Hans zu einem anderen Zeitpunkt auch stechende Schmerzen im kleinen Finger seiner rechten Hand verspürt, dass sich aber gleichzeitig die physikalische Realisierung dieser Schmerzen signifikant vom Zustand N unterscheidet. Zudem ist natürlich auch vorstellbar, dass das Empfinden von stechenden Schmerzen im kleinen Finger der rechten Hand bei einer anderen Person, etwa Hanna, von einem neuronalen Zustand begleitet wird, der entscheidend von N divergiert. Daher kommt Yablo zu folgendem Fazit, in dem er gleichzeitig seinen Lösungsvorschlag für das Exklusionsproblem präsentiert:

So I find no fault with dualism, or with the associated picture of mental phenomena as necessitated by physical phenomena which they are possible without. Rather than objecting, in fact, to the asymmetric necessitation picture, I propose to go it one better. Traditionally, the paradigm of one-way necessitation was the relation of *determinate* and *determinable* (...). What if mental phenomena are determinables of physical phenomena in something like the traditional sense? Then since a determinate cannot preempt its own determinable, mental events and properties lose nothing in causal relevance to their physical bases (...). If anything, it is the other way around. Overladen as they frequently are with physical details far beyond the effect's causal requirements, it is the *physical* phenomena which are liable to disqualification on ground of superfluity (...).²⁶⁷

Das Verhältnis zwischen mentalen und physikalischen Eigenschaften ist für Yablo ein klassisches Beispiel für eine Relation asymmetrischen Notwendigmachens. Traditionell wird eine solche Beziehung als Determinaten-Determinablen-Relation widergegeben. Dementsprechend lautet Yablos Lösungsvorschlag für das Exklusionsproblem: mentale Eigenschaften müssen

²⁶⁶ vgl. Yablo, Stephen: „*Mental Causation*“ in: *The Philosophical Review*, Vol. 101, 1992, S. 248ff.

²⁶⁷ Yablo, Stephen: „*Mental Causation*“ in: *The Philosophical Review*, Vol. 101, 1992, S. 250.

als Determinablen physikalischer Eigenschaften, ihrer Determinaten, verstanden werden. Da Determinaten ihre Determinablen nicht der kausalen Kraft berauben, verlieren mentale Eigenschaften hinsichtlich ihrer kausalen Relevanz auch nichts an physikalische Eigenschaften. Im Gegenteil, es sind vielmehr die hochspezifischen physikalischen Eigenschaften, die Informationen enthalten, die die kausalen Erfordernisse bei weitem überschreiten und daher in Kausalerklärungen eher vernachlässigt werden dürfen.

Doch was sind Determinaten und Determinablen eigentlich? Diese Frage beantwortet Yablo anhand von folgendem Beispiel: „(...) some properties stand to others as *determinate* to *determinable* – for example, *crimson* is a determinate of the determinable *red*, *red* is a determinate of *colored*, and so on. Since the distinction is relative, one does better to speak of a determination *relation* among properties (...)“²⁶⁸ Determinaten und Determinablen kennzeichnen also ein Verhältnis von Eigenschaften zueinander: scharlachrot ist beispielsweise eine Determinate der Determinable rot, rot wiederum ist selbst eine Determinate der Determinable farbig, usw. Dieses Beispiel enthüllt auch den relativistischen Charakter der Determinaten-Determinablen-Beziehung: im Beispiel ist ‚rot‘ eine Determinable der Determinate ‚scharlachrot‘, gleichzeitig aber auch eine Determinate der Determinable ‚farbig‘. Formal gibt Yablo die Determinierungsrelation folgendermaßen wieder:

(Δ) *P* determines *Q* iff: for a thing to be *P* is for it to be *Q*, not *simpliciter*, but in a specific way.

Except for the ‚not *simpliciter* ...‘, (Δ) would describe identity (...). Metaphysically, the central idea is that

- (Δ) *P* determines *Q* (*P* > *Q*) only if:
- (i) necessarily, for all *x*, if *x* has *P* then *x* has *Q*; and
 - (ii) possibly, for some *x*, *x* has *Q* but lacks *P*.²⁶⁹

Eine Eigenschaft *P* determiniert also eine Eigenschaft *Q* genau dann, wenn es für etwas bedeutet *P* zu sein, dies auch gleichzeitig bedeutet *Q* zu sein, jedoch nicht *simpliciter*, sondern auf eine bestimmte Art und Weise. Durch diese spezielle Art des Gegebenseins unterscheidet sich die Determinierungsrelation auch von einer Identitätsbeziehung. Modal ausgedrückt determiniert eine Eigenschaft *P* eine Eigenschaft *Q*, wenn notwendigerweise für alle Dinge *x* gilt, wenn *x* *P* hat, dann hat *x* *Q* und wenn gleichzeitig die Möglichkeit besteht, dass es einige Dinge *x* gibt, die *Q* sind und zudem nicht *P* sind.

²⁶⁸ Yablo, Stephen: „*Mental Causation*“ in: *The Philosophical Review*, Vol. 101, 1992, S. 252.

²⁶⁹ Yablo, Stephen: „*Mental Causation*“ in: *The Philosophical Review*, Vol. 101, 1992, S. 252.

Im Hinblick auf mentale und physikalische Eigenschaften geht Yablo davon aus, dass sie genau in einer solchen Determinierungsbeziehung zueinander stehen. Nämlich, dass physikalische Eigenschaften mentale Eigenschaften determinieren. Folglich gelangt er diesbezüglich zur Schlussfolgerung: „(...) mental/physical relations are a species of determinable/determinate relations.“²⁷⁰

Anschließend analysiert Yablo das Exklusionsprinzip und zwar insbesondere im Hinblick auf dessen Applikation auf Eigenschaften. In diesem Rahmen stellt Yablo folgende zwei Beispiele vor:

Imagine a pigeon, Sophie, conditioned to peck at red to the exclusion of other colors; a red triangle is presented, and Sophie pecks. Most people would say that the redness was causally relevant to her pecking, even that this was a paradigm case of causal relevance. But wait! I forgot to mention that the triangle in question was a specific shade of red: scarlet. Assuming that the scarlet was causally sufficient for the pecking, we can conclude by the exclusion principle that every *other* property was irrelevant. Apparently, then, the redness, although it looked to be *precisely* what Sophie was responding to, makes in reality no causal contribution whatever. Another example concerns properties of events. Suppose that the structures in a certain region, though built to withstand lesser earthquakes, are in the event of a *violent* earthquake – one registering over five on the Richter scale – causally guaranteed to fall. When one unexpectedly hits, and the buildings collapse, one property of the earthquake that seems relevant to their doing so is that it was violent. Or so you might think, until I add that this particular earthquake was *barely* violent (its Richter magnitude was over five but less than six). What with the earthquake's *bare* violence being *already* causally sufficient for the effect, that it was *violent* made no causal difference.²⁷¹

Das erste Beispiel handelt von einer Taube namens Sophie, die so konditioniert wurde, dass sie bei der Präsentation von mehreren, verschiedenfarbigen Objekten immer das rote Objekt ausfindig machen kann, indem sie auf es pickt. Wird Sophie nun etwa ein rotes Dreieck zusammen mit beispielsweise einem grünen Kreis und ein blauen Quadrat vorgelegt, dann pickt sie auf das rote Dreieck. Intuitiv würden die meisten Menschen wohl sagen, dass die Röte des Dreiecks kausal relevant für Sophies Picken war. Man würde geradezu davon ausgehen, dass es sich hier um ein Paradebeispiel kausaler Relevanz handelt. Allerdings hatte das Dreieck einen bestimmten Rotton – es war nämlich scharlachrot. Nimmt man nun an, dass das Scharlachrote kausal hinreichend für Sophies Picken war, dann kann man mit dem Exklusionsprinzip schließen, dass jegliche andere Eigenschaft kausal irrelevant war. Folglich leistet auch die Röte des Dreiecks – entgegen aller primärer Erwartung, dass es genau dies war, worauf Sophie reagiert hat – keinen kausalen Beitrag zur Verursachung des Pickens. Zu einem ähnlich kontraintuitiven Ergebnis gelangt auch das zweite von Yablo eingeführte Exempel. In einer bestimmten Region sind alle Gebäude so gebaut, dass sie leichten Erdbeben standhalten. Ein starkes Erdbeben, d.h. ein Beben mit einer Stärke von mindestens fünf auf der Richterskala, wäre jedoch kausal hinreichend für das Einstürzen der Gebäude in dem betroffenen Gebiet. Trifft nun ein solches starkes Erdbeben die Region fallen die Gebäude in sich zusammen.

²⁷⁰ Yablo, Stephen: „*Mental Causation*“ in: *The Philosophical Review*, Vol. 101, 1992, S. 256.

²⁷¹ Yablo, Stephen: „*Mental Causation*“ in: *The Philosophical Review*, Vol. 101, 1992, S. 257f.

Eine wesentliche Eigenschaft, die für das Einstürzen der Gebäude kausal verantwortlich zu sein scheint, besteht eben darin, dass es ein starkes Erdbeben war. Nun könnte man aber zusätzlich angeben, dass es ein gerade-noch-starkes Erdbeben war – ‚gerade-noch-stark‘ bedeutet in diesem Zusammenhang, dass seine Stärke einen Wert zwischen fünf und sechs auf der Richterskala hatte. Ist aber die Gerade-noch-Stärke des Erdbebens bereits kausal hinreichend für das Kollabieren der Häuser, dann scheint es kausal keinen Unterschied zu machen, ob das Erdbeben auch ein starkes Erdbeben war oder nicht. Insofern ist auch die Eigenschaft, dass das Erdbeben ein starkes Erdbeben war nicht, wie ursprünglich angenommen, kausal relevant für das Einstürzen der Häuser in der betroffenen Region. Auch dieses Ergebnis widerspricht der Intuition. In der Tat ist es sehr überraschend, dass sowohl die Röte des Dreiecks als auch die Stärke des Erdbebens nicht, wie zunächst vermutet, die kausal relevanten Eigenschaften in den jeweiligen Kausalbeziehungen darstellen. Daher stellt Yablo umgehend folgende Frage:

(...) if even paradigm cases of causal relevance fail the exclusion test, what passes it? Not much, it turns out. Almost whenever a property *Q* is *prima facie* relevant to an effect, a causally sufficient determination *Q'* of *Q* can be found to expose it as irrelevant after all. Applying the argument to *Q'*, *Q''*, etc. in turn, it appears that only ultimate determinates – properties unamenable to further determination – can hope to retain their causal standing.

Or, on second thought, maybe not them either. Not everything about a cause contributes to its effect; and even where a property does contribute, it need not do so in all aspects. From the examples it is clear that such irrelevancies do indeed creep in, as we pass from determinable to determinate (e.g., registering less than six); and if the determination process is continued *ad finem*, they may be expected to accumulate significantly. So any ultimate determinate seems likely to incorporate causally extraneous detail. But then, abstracting some or all of this detail away should leave a determinable which, since it falls short of the original only in irrelevant respects, is no less sufficient for the effect. By the exclusion principle, this robs even ultimate determinates of their causal powers. And now it begins to look as though no property ever makes any causal difference.

At least as it applies to properties, then, the exclusion principle is badly overdrawn.²⁷²

Wenn also selbst Paradebeispiele kausaler Relevanz, wie die Röte des Dreiecks und die Stärke des Erdbebens in obigen Beispielen, dem Exklusionsprinzip zum Opfer fallen, scheint einem ‚Wettlauf nach unten‘ damit Tür und Tor geöffnet. Denn für jede Eigenschaft *Q*, die zunächst als kausal relevant für eine Wirkung erscheint, kann eine kausal hinreichende Determinierung *Q'* gefunden werden, die *Q* als kausal irrelevant entlarvt. Wendet man dasselbe Prinzip wiederum auf *Q'*, *Q''*, usw. an gerät man in den Sog einer Abwärtsspirale, die erst bei so genannten ultimativen Determinaten Halt macht. Diese ultimativen Determinaten sind Eigenschaften, die sich einer weiteren Determinierung gegenüber verschließen. Aber selbst die ultimativen Determinaten scheinen bei genauerer Betrachtung ihre kausale Bedeutung (zumindest teilweise) einzubüßen. Denn nicht alle Eigenschaften einer Ursache tragen in gleicher Weise zur Wirkung bei und selbst wenn eine Eigenschaft einen kausalen Beitrag leistet, so tut sie dies nicht in allen Aspekten. An den obigen Beispielen dürfte klar geworden sein,

²⁷² Yablo, Stephen: „*Mental Causation*“ in: *The Philosophical Review*, Vol. 101, 1992, S. 258f.

dass sich auf dem Weg von Determinable zu Determinate langsam zahlreiche solcher Belanglosigkeiten (wie z.B. ‚hat einen Wert unter sechs auf der Richterskala‘) einschleichen. Dies natürlich umso mehr je weiter der Determinierungsprozess vorangetrieben wird. Daher scheinen ultimative Determinaten einen langen Katalog an kausal irrelevanten Details zu beinhalten. Lässt man nun aber die für die Verursachung überflüssigen Details weg, gelangt man zu einer Determinable, die genauso kausal hinreichend für die Wirkung ist wie die ursprüngliche ultimative Determinate – schließlich wurden ausschließlich für die Verursachung irrelevante Einzelheiten weggekürzt. Dementsprechend verlieren auf diese Weise auch ultimative Determinaten ihre kausale Kraft und es scheint als würde keine Eigenschaft überhaupt jemals einen kausalen Unterschied machen. Daraus folgert Yablo, dass das Exklusionsprinzip – zumindest im Hinblick auf Eigenschaften – stark überzogen ist. Selbst eine Überarbeitung des Exklusionsprinzips hätte keine Auswirkungen auf Determinaten und Determinablen, denn

(...) we know that the corrected principle does not apply to determinates and their determinables – for we know that they are not causal rivals. This kind of position is of course familiar from other contexts. Take for example the claim that a space completely filled by one object can contain no other. Then are even the object's *parts* crowded out? No. In this competition wholes and parts are not on opposing teams; hence any principle that puts them there needs rethinking. Likewise any credible reconstruction of the exclusion principle must respect the truism that determinates do not contend with their determinables for causal influence.²⁷³

Im Grunde geht das Exklusionsprinzip im Falle mentaler Verursachung also in die Irre, weil es Eigenschaften, die zueinander in einer Determinaten-Determinablen-Relation stehen, wie kausale Wettbewerber behandelt. Ähnlich wie Objekte, die in einem Teil-Ganzes-Verhältnis zueinander stehen, berauben die einen die anderen jedoch nicht ihrer kausalen Relevanz. Nimmt man beispielsweise an, dass ein Raum, der komplett von einem Gegenstand ausgefüllt wird, gleichzeitig keinen anderen Gegenstand beherbergen kann, so gilt dies jedoch offensichtlich nicht für Teile des Gegenstands. Jedes Prinzip, das ein solches Konkurrenzverhältnis postuliert, müsse folglich auch überdacht werden, so Yablo. Analog bedarf auch das Exklusionsprinzip im Hinblick auf mentale Eigenschaften einer neuen Evaluation. Dementsprechend gelangt Yablo in Anbetracht mentaler Verursachung zu folgendem Fazit:

With the exclusion principle neutralized, the application to mental causation is anticlimactic. As a rule, determinates are tolerant, indeed supportive, of the causal aspirations of their determinables. Why should it be different, if the determinate is physical and the determinable is mental? Inferring the causal irrelevance of, say, my *dizziness*, from the causal sufficiency of its physical basis, is not appreciably better than rejecting the redness as irrelevant on the ground that all the causal work is accomplished already by its determinate scarlet. Or, if someone thinks it *is* better, then she owes us an explanation of what the metaphysically important difference is between the cases. That there is a conceptual difference is granted, but it is not to the point; there is no conceptual entailment either from the tea's micromechanical condition to its high temperature, yet this occasions little skepticism

²⁷³ Yablo, Stephen: „*Mental Causation*“ in: *The Philosophical Review*, Vol. 101, 1992, S. 259.

about the role of the tea's temperature in its burning my tongue. If there is a metaphysical difference, then someone should say what it is, and why it matters to causation.²⁷⁴

Da Determinablen nicht mit ihren Determinaten um kausale Relevanz konkurrieren, sind diese auch nicht durch das Exklusionsprinzip von kausaler Wirkungslosigkeit bedroht. Im Gegenteil, Determinaten unterstützen sogar die kausalen Bestrebungen ihrer Determinablen. Die kausale Belanglosigkeit meines Schwindelgefühls daraus zu schließen, dass dessen physikalische Basis hinreichend für seine Verursachung ist, wäre laut Yablo in etwa genauso, wie die kausale Wirksamkeit der Röte in obigem Beispiel zu bestreiten, weil die gesamte kausale Arbeit schon vom Scharlachrot geleistet wird. Sollte jemand diese Analogie anzweifeln, so müsste er auch den metaphysisch relevanten Unterschied zwischen beiden Fällen herausarbeiten, so Yablo.

Am Ende seines Aufsatzes führt Yablo folgendes Beispiel an, um seine Position nochmals zu untermauern:

I arrive on your doorstep and, rather than knocking, decide to press the buzzer. Epiphenomenalist neuroscientists are monitoring my brain activity from a remote location, and an event *e* in their neurometer indicates my neural condition to be such and such. Now, like any mental event, my decision *m* has a physical determination *p*, and the question arises to which of these the neurometer reading *e* is due. The scientist reasons as follows: Because the neurometer is keyed to the precise condition of his brain, *e* would not have occurred if the decision had been taken in a different neural way, in particular if it had occurred in *p*'s absence. So *m* was not enough for *e*; *p* on the other hand looks roughly proportional to *e* and so has the better claim to cause it. Another triumph for epiphenomenalism!²⁷⁵

In Yablos Beispiel kommt eine Person an einer Haustüre an und entscheidet sich die Klingel zu betätigen anstatt anzuklopfen. Währenddessen messen Neurowissenschaftler aus einer gewissen Entfernung die Hirnaktivität der Person. Dabei zeigt das Ereignis *e* auf deren Neurometer an, dass sich die Person in so-und-so-einem neuronalen Zustand befindet. Wie jedes andere mentale Ereignis, so hat auch *m*, die Entscheidung der Person zu klingeln, eine physikalische Determinierung *p*. Nun stellt sich die Frage auf welches Ereignis *e* zurückzuführen ist – auf *m* oder auf *p*? Der Argumentationsgang eines Neurowissenschaftlers in Hinblick auf diese Frage lässt sich in etwa so rekonstruieren: Die Anzeige auf dem Neurometer hängt unmittelbar mit dem Zustand des Gehirns der klingelnden Person zusammen – *e* hätte sich also nicht ereignet, wenn die Entscheidung auf eine andere Art und Weise neuronal realisiert worden wäre – genauer gesagt, wenn sie in Abwesenheit von *p* eingetreten wäre. Folglich war das Eintreten von *m* nicht ausreichend für das Sich Ereignen von *e*. Andererseits sieht es so aus als wäre *p* ungefähr proportional zu *e*. Daher hat es auch einen größeren Anspruch als *m* die Ursache von *e* zu sein. Laut Yablo ist der Gedankengang des Neurowissenschaftlers richtig

²⁷⁴ Yablo, Stephen: „*Mental Causation*“ in: *The Philosophical Review*, Vol. 101, 1992, S. 260.

²⁷⁵ Yablo, Stephen: „*Mental Causation*“ in: *The Philosophical Review*, Vol. 101, 1992, S. 277f.

bis auf den letzten Schritt in der Argumentationskette. Um seine Überlegung zu untermauern hinterfragt Yablo zunächst unsere alltägliche explanatorische Praxis: wann schreiben wir mentalen Ursachen Wirkungen zu? In der Regel nur dann, wenn wir annehmen, dass die Wirkung relativ unempfindlich gegenüber den Details der physikalischen Implementierung von *m* ist.²⁷⁶ Auf das Türglockenbeispiel übertragen bedeutet dies: „Having decided to push the button, I do so, and the doorbell rings. Most people would say, and I agree, that my decision had the ringing as one of its effects. Of course, the decision had a physical determination *p*; but, most people would also say, and I agree again, that it would still have been succeeded by the ringing, if it had occurred in a different physical way, that is, if its physical determination had not been *p* but some other physical event. And this is just to say that *p* was not *required* for the effect.“²⁷⁷ Nachdem sich die Person also entschieden hatte zu klingeln, drückte sie den Knopf, worauf die Türglocke läutete. Gemeinhin würde man nun sagen, dass das Läuten der Türglocke eine der Wirkungen der Entscheidung zu klingeln war. Gewiss hatte diese Entscheidung auch eine physikalische Determinierung *p*. Allerdings wäre der Entscheidung auch dann ein Glockenläuten gefolgt, wenn dessen physikalische Implementation eine andere gewesen wäre. Das heißt aber nichts anderes als, dass *p* nicht nötig war für das Eintreten der Wirkung. Die physikalische Realisierung *p* der Entscheidung zu klingeln spielt folglich eine untergeordnete Rolle für das Eintreten der Wirkung (i.e. des Läutens an der Türglocke). Deshalb kann *p* wohl auch kaum als ein adäquater Kandidat für die Ursache des Türklingelns gehandelt werden.

Yablos eigene Rekonstruktion der kausalen Verhältnisse, in dem von ihm vorgestellten Beispiel sieht folgendermaßen aus (wobei *r* für das Läuten der Türglocke steht):

- (i) *m* is a counterexample to *r*'s requiring *p* (for *r* would still have occurred, if *m* had occurred without *p*);
- (ii) *p* is not proportional to *r* (since *r* does not require it);
- (iii) *p* does not cause *r* (since it is not proportional to *r*);
- (iv) *p* is not a counterexample to *m*'s enoughness for *r* (it could be a counterexample only if *r* required it);
- (v) *p* is not a counterexample to *m*'s proportionality with *r* (by inspection of the remaining conditions);
- (vi) *p* poses no evident threat to the hypothesis that *m* caused *r*.

Here are the beginnings, at least, of a story wherein a mental event emerges as better qualified than its physical basis for the role of cause. I believe that this kind of story is enacted virtually wherever common sense finds mental causation.²⁷⁸

Yablo zufolge stellt also *m* ein Gegenbeispiel dazu dar, dass *p* für *r* nötig ist. Denn hätte sich *m* ereignet und *p* nicht, so wäre *r* trotzdem eingetreten (i). Natürlich hätte *m* in diesem Fall

²⁷⁶ vgl. Yablo, Stephen: „Mental Causation“ in: *The Philosophical Review*, Vol. 101, 1992, S. 278.

²⁷⁷ Yablo, Stephen: „Mental Causation“ in: *The Philosophical Review*, Vol. 101, 1992, S. 278.

²⁷⁸ Yablo, Stephen: „Mental Causation“ in: *The Philosophical Review*, Vol. 101, 1992, S. 278f.

eine andere physikalische Realisierung als p besessen. Da p somit nicht nötig für r ist, kann p auch nicht proportional zu r sein (ii). Dies wiederum hat zur Folge, dass p aufgrund seiner mangelnden Proportionalität zu r nicht als dessen Ursache in Frage kommt (iii). Schließlich müssen einer weit verbreiteten Annahme zufolge Ursache und Wirkung proportional zueinander sein. p wiederum kann laut Yablo kein Gegenbeispiel dafür sein, dass m ausreichend für r ist. Dies wäre nämlich nur dann möglich, wenn p für r nötig wäre (iv). Wäre dies der Fall, könnte r nämlich nicht eintreten, wenn p nicht vorher eingetreten wäre. Da aber r durchaus selbst dann noch eintreten kann, wenn p nicht eintritt (wenn m also eine andere physikalische Implementierung besitzt), ist es mitunter auch möglich, dass m für das Eintreten von r ausreichend ist. Insofern kann p auch kein Gegenbeweis dafür sein, dass m proportional zu r ist (v). Denn auch die Proportionalität m 's r gegenüber würde selbstverständlich unterminiert werden, wenn p für r nötig wäre. Dies ist jedoch nicht der Fall und dadurch kann p kein Gegenargument zu der These, dass m proportional zu r ist, sein. Daraus leitet Yablo seine finale Konklusion ab, dass p die Annahme m sei die Ursache von r keineswegs untergräbt (vi). Im Gegenteil, dies stütze sogar die These, dass m eigentlich der bessere Ursachenkandidat für r ist als p . Außerdem decke sich diese Sichtweise auch mit dem common sense Verständnis von mentaler Verursachung. Schließlich werde in der alltäglichen explanatorischen Praxis von Verhalten auch z.B. auf Gründe verwiesen und nicht etwa auf Hirnzustände.

Kausale Wirksamkeit und kausale Relevanz

In ihrem, im Jahr 1990 in der Zeitschrift *Analysis* publizierten, Aufsatz *Program Explanation: A General Perspective* beziehen sich Frank Jackson und Philip Pettit zwar nicht unmittelbar auf das Exklusionsproblem, aber die dort von ihnen vorgeschlagene ‚Programmerklärung‘ kann als eine geradezu paradigmatische Anwendung auf die vom Exklusionsargument aufgeworfene Problematik verstanden werden. Ganz allgemein geht es Jackson und Pettit um die Rolle von Eigenschaften in kausalen Erklärungen. Genauer gesagt behandeln sie ein Problem, das bei der gleichzeitigen Billigung von vier intuitiv plausiblen Annahmen entsteht und skizzieren anschließend eine Lösung für dieses Problem. Die ersten drei Annahmen über die Rolle von Eigenschaften in Kausalerklärungen bilden Jackson und Pettit zufolge den Hintergrund für das Problem und werden von ihnen folgendermaßen eingeführt:

1. A causal explanation of something must direct us to a causal relevant property as opposed to a causally irrelevant property of the factor it identifies as explanatory: a property relevant to the causal production of the effect explained.

2. One way in which properties are causally relevant is by being causally efficacious. A causally efficacious property with regard to an effect is a property in virtue of whose instantiation, at least in part, the effect occurs; the instance of the property helps to produce the effect and does so because it is an instance of that property.
3. A property F is not causally efficacious in the production of an effect *e* if these three conditions are fulfilled together.
 - (i) there is a distinct property G such that F is efficacious in the production of *e* only if G is efficacious in its production;
 - (ii) the F-instance does not help to produce the G-instance in the sense in which the G-instance, if G is efficacious, helps to produce *e*; they are not sequential causal factors;
 - (iii) the F-instance does not combine with the G-instance, directly or via further effects, to help in the same sense to produce *e* (nor of course, *vice versa*): they are not coordinate causal factors.²⁷⁹

Die erste Annahme ist laut Jackson und Pettit kaum weiter erklärungsbedürftig. Denn es ist klar, dass die kausale Erklärung eines Ereignisses zu den Eigenschaften hinführen muss, die kausal relevant für das Eintreten der entsprechenden Wirkung sind. Dabei spiele es keine Rolle auf welche Art und Weise die Ursachen genau individuiert oder identifiziert werden. Die zweite Hypothese über kausal relevante Eigenschaften ist ähnlich offensichtlich. Eine Möglichkeit für Eigenschaften kausal relevant zu sein, ist es, kausal wirksam²⁸⁰ zu sein. Eine, im Hinblick auf eine bestimmte Wirkung kausal wirksame, Eigenschaft ist definiert als eine Eigenschaft, aufgrund derer Instanziierung die Wirkung (zumindest zum Teil) eintritt. Das Haben der Eigenschaft ist wesentlich daran beteiligt die Wirkung hervorzubringen und zwar gerade weil es eine Instanziierung eben dieser Eigenschaft ist. Ferner merken Jackson und Pettit an, dass ‚Wirksamkeit‘ in diesem Zusammenhang nicht unbedingt als nicht reduzierbarer Bestandteil der Welt verstanden werden muss. Vielmehr kann ‚Wirksamkeit‘ auch dekonstruiert werden, etwa im Sinne von Kausalgesetzen. Schließlich gilt es auch ‚kausale Wirksamkeit‘ penibel von ‚instrumenteller Effektivität‘ zu unterscheiden. Eine Eigenschaft ist genau dann instrumentell effektiv hinsichtlich einer bestimmten Wirkung, wenn es eine vielsprechende Strategie für die Herbeiführung der Wirkung ist, diese Eigenschaft zu verwirklichen. Es ist offenkundig, dass instrumentelle Effektivität in diesem Sinne kausale Wirksamkeit nicht miteinschießt – denn die Wirkung tritt in diesem Fall nicht deshalb ein, weil die Eigenschaft instanziiert wurde. Die dritte Annahme besagt, dass eine Eigenschaft F nicht kausal wirksam für das Herbeiführen einer Wirkung *e* ist, wenn folgende drei Bedingungen gleichzeitig erfüllt sind. Erstens muss es eine Eigenschaft G geben, die von F verschieden ist, und zugleich muss gelten: F ist nur dann wirksam für das Eintreten von *e*, wenn auch G wirksam an dessen Herbeiführung beteiligt ist. Zweitens ist die Instanziierung von F nicht am Hervorbringen der Instanziierung von G beteiligt, in dem Sinn wie G, wenn es kausal wirksam ist, am Bewirken

²⁷⁹ Jackson, Frank & Pettit, Philip: „*Program Explanation – A General Perspective*“, in: *Analysis*, Vol. 50, 1990, S. 108.

²⁸⁰ Jackson und Pettit verwenden im englischsprachigen Original die Begriffe ‚causally relevant‘ und ‚causally efficacious‘, diese werden von mir als ‚kausal relevant‘ und ‚kausal wirksam‘ wiedergegeben.

von *e* Anteil hat. F und G sind also keine kausale Abfolge derart, dass zuerst F G und hinterher G *e* verursacht. Drittens ist es auch keine mittel- oder unmittelbare Verbindung aus den Instanziierungen von F und G, die *e* bewirkt. D.h. F und G sind keine gleichgestellten kausalen Faktoren bei der Verursachung von *e*. Die gesamte dritte Annahme ist nach Jackson und Pettit intuitiv nicht in derselben Weise einleuchtend wie die ersten beiden und bedarf deshalb weiterer Erläuterung. Eine Möglichkeit für F nur dann kausal wirksam zu sein, wenn G kausal wirksam ist, besteht darin, dass die Instanziierungen von F und G eine verhältnismäßig weitere bzw. nähere Ursache von *e* sind. Dieser Fall wäre zwar durchaus mit der Annahme 3(i) vereinbar, würde aber aufgrund des sequentiellen Charakters von F und G durch 3(ii) ausgeschlossen. 3(iii) wiederum, dient dazu 3(i) so einzuschränken, dass Fälle, in denen F und G beides notwendige Teile eines komplexes Kausalprozesses sind, nicht als Vorkommnisse gelten, bei denen die Instanziierung von F nur dann kausal wirksam ist, wenn es die Instanziierung von G ist.²⁸¹ Daher bilanzieren Jackson und Pettit: „Notice that for all that 3 says, the instances of the distinct properties F and G may be identical. If G is efficacious in such a case, being a property in virtue of whose instantiation *e* occurs, still that will not make F efficacious: the instance of F will help to produce *e* but not because it is an instance of F; instead it will do so because it is an instance of G.“²⁸² Nach 3. können also die Instanziierungen der unterschiedlichen Eigenschaften F und G durchaus identisch sein. Wenn G nämlich in einem solchen Fall kausal wirksam ist, d.h. wenn *e* eintritt, weil G instanziiert wurde, heißt das nicht, dass gleichzeitig auch F kausal wirksam ist. Schließlich bedeutet dies zwar, dass F am Herbeiführen von *e* beteiligt war, aber nicht etwa weil es eine Instanziierung der Eigenschaft F war, sondern weil es zugleich auch eine Instanziierung der Eigenschaft G war.

Als erstes Beispiel für eine Situation, in der 3(i), 3(ii) und 3(iii) gemeinsam wahr sind, führen Jackson und Pettit folgendes Szenario an:

(A) A fragile glass is struck and breaks. Why did it break? First answer: because of its fragility. Second answer: because of the particular molecular structure of the glass. The property of fragility was efficacious in producing the breaking only if the molecular structural property was efficacious: hence 3(i). But the fragility did not help to produce the molecular structure in the way in which the structure, if it was efficacious, helped to produce the breaking. There was no time-lag between the exercise of the efficacy, if it was efficacious, by the disposition and the exercise of the efficacy, if it was efficacious, by the structure. Hence 3(ii). Nor did the fragility combine with the structure, in the manner of a coordinate factor, to help in the same sense to produce *e*. Full information about the structure, the trigger and the relevant laws would enable one to predict *e*; fragility would not need to be taken into account as a coordinate factor. Hence 3(iii).²⁸³

²⁸¹ vgl. Jackson, Frank & Pettit, Philip: „Program Explanation – A General Perspective“, in: *Analysis*, Vol. 50, 1990, S. 108f.

²⁸² Jackson, Frank & Pettit, Philip: „Program Explanation – A General Perspective“, in: *Analysis*, Vol. 50, 1990, S. 109.

²⁸³ Jackson, Frank & Pettit, Philip: „Program Explanation – A General Perspective“, in: *Analysis*, Vol. 50, 1990, S. 109.

Im ersten Beispiel von Jackson und Pettit prallt ein zerbrechliches Glas irgendwo auf und zerbricht. Auf die Frage warum das Glas zerbrochen ist, scheint es zwei (kausale) Erklärungen zu geben. Die erste Antwort besagt, weil das Glas zerbrechlich war, die zweite, weil das Glas eine bestimmte molekulare Struktur auswies. Hierbei ist die Eigenschaft der Zerbrechlichkeit jedoch nur dann kausal wirksam am Zerbrechen des Glases beteiligt, wenn auch die Eigenschaft so-und-so-eine-molekulare-Struktur-zu-haben kausal wirksam ist. Aus diesem Grund gilt 3(i) in dieser Konstellation. Zudem hatte die Zerbrechlichkeit nicht in der gleichen Weise Anteil am Herbeiführen des Zersplittern des Glases wie die molekulare Struktur des Glases, wenn diese kausal wirksam war. Es gab keinen zeitlichen Abstand zwischen der kausalen Wirksamkeit der dispositionalen Eigenschaft zerbrechlich zu sein und der kausalen Wirksamkeit der molekularen Struktur. Daher gilt 3(ii). Schließlich haben auch nicht die Zerbrechlichkeit und die Strukturbeschaffenheit des Glases gemeinsam, sozusagen als gleichberechtigte Ursachen, das Zerbrechen des Glases bewirkt. Denn ein Wissen um die molekulare Struktur des Glases, den Auslöser des Zerbrechens und die wirkenden Kausalgesetze hätte bereits vollkommen ausgereicht, um *e* präzise vorherzusagen. Ein Referieren auf so etwas wie die Zerbrechlichkeit des Glases als (der Struktur gleichgestellter) Kausalfaktor wäre dazu nicht nötig gewesen. Somit trifft auf dieses Beispiel schlussendlich auch 3(iii) zu. Als zweites Beispiel führen Jackson und Pettit folgende Konstellation ein:

(B) I try and fail to fit a square peg in a round hole of diameter equal to the side of the square. Why did it not go through? First answer: because of the squareness of the peg. Second answer: because of the impenetrability of this overlapping part of the peg. The property of squareness was efficacious only if the overlap-cum-impenetrability-property was efficacious: hence 3(i) is true. But 3 (ii) is also true, for the squareness did not help to produce the overlap-cum-impenetrability in the way in which it, if it was efficacious, helped to produce the blocking of the peg: there was no time-lag of the sort that such an influence would seem to require. And 3(iii) is also true, for the squareness did not combine with the overlap-cum-impenetrability to help in the same sense to produce the blocking; one could have predicted the blocking without reference to squareness. As we might put it, the overlap-cum-impenetrability did not need any extra help from the squareness to produce the blocking.²⁸⁴

In diesem Szenario wird versucht einen quadratischen Pfahl durch eine runde Öffnung zu schieben, wobei der Durchmesser der Öffnung so groß ist wie eine Seitenlänge der quadratischen Grundform des Pflocks. Es ist offensichtlich, dass es nicht gelingen kann den Pfahl durch die Öffnung zu schieben. Warum? Die erste mögliche Antwort lautet: „Wegen der quadratischen Grundform des Pflocks.“ Die zweite mögliche Replik könnte etwa sein: „Wegen der Undurchdringlich- bzw. Nichtdurchlässigkeit des überstehenden Pfahlstücks.“ Nun war die Eigenschaft eine-quadratische-Grundform-zu-haben nur dann kausal wirksam, wenn es auch die Überlappen-plus-Undurchdringlichkeit-Eigenschaft war. Daraus leitet sich die Gül-

²⁸⁴ Jackson, Frank & Pettit, Philip: „*Program Explanation – A General Perspective*“, in: *Analysis*, Vol. 50, 1990, S. 110.

tigkeit von 3(i) ab. Allerdings muss auch 3(ii) gelten, denn die Eigenschaft eine-quadratische-Grundform-zu-haben war nicht in gleicher Weise an der Verursachung der Überlappen-plus-Undurchdringlichkeit-Eigenschaft beteiligt, wie es letztere an der Unmöglichkeit den Pfahls durch die Öffnung zu führen war, wenn sie tatsächlich kausal wirksam war. Schließlich gab es keinen zeitlichen Abstand zwischen den Instanziierungen dieser beiden Eigenschaften, so wie es zu erwarten gewesen wäre, wenn eine Kausalrelationen zwischen beiden bestünde. Da die Eigenschaft eine-quadratische-Grundform-zu-haben und die Überlappen-plus-Undurchdringlichkeit-Eigenschaft nicht sozusagen im Konzert die Unmöglichkeit den Pflock durch die Öffnung zu schieben verursacht haben, trifft auch 3(iii) zu. Insgesamt wäre eine Vorhersage des Nicht-Einführbarkeit des Pfahls auch ohne Bezug auf dessen Eigenschaft eine-quadratische-Grundform-zu-haben möglich gewesen. Mit anderen Worten war die Überlappen-plus-Undurchdringlichkeit-Eigenschaft bei der Verursachung des Nicht-Penetrierbarkeit der Öffnung durch den Pflock nicht auf die ‚Schützenhilfe‘ der Eigenschaft eine-quadratische-Grundform-zu-haben angewiesen. Hiermit trifft auch in diesem Beispiel 3(iii) zu. Das dritte und letzte Exempel von Jackson und Pettit ist das Folgende:

(C) The water in a closed glass container reaches boiling temperature – the mean molecular motion is at such and such a level – and the container cracks. Why did it crack? First answer: because of the temperature of the water. Second answer, in simplified form: because of the momentum of such and such a molecule (group of molecules) in striking such and such a molecular bond in the container surface. (We are supposing that the case is one where the container breaks because of the internal pressure, not because of the temperature gradient between the water and the container.) The temperature-property was efficacious only if the momentum-property was efficacious: hence 3(i). But the temperature of the water – an aggregate statistic – did not help to produce the momentum of the molecule in the way in which it, if efficacious, helped to produce the cracking: hence 3(ii). And neither did the temperature combine with the momentum to help in the same sense to produce the cracking: one could have predicted the cracking just from full information about the molecule and the relevant laws. Hence 3(iii).²⁸⁵

In einem verschlossenen, gläsernen Behältnis wird Wasser solange erhitzt, bis es zu kochen beginnt woraufhin das Gefäß einen Sprung bekommt. Auf die Frage warum der Behälter ge-
borsten ist, gibt es wiederum zwei Antworten. Zum Einen könnte man antworten, dies sei aufgrund der Temperatur des Wassers geschehen. Zum Anderen wäre auch diese (etwas vereinfachte) Antwort denkbar: Wegen des Impulses diesen-oder-jenen Moleküls (bzw. dieser-oder-jener Gruppe von Molekülen) beim Auftreffen auf eine Molekularbindung der Oberfläche des Gefäßes. Bei der Antwort wird stillschweigend vorausgesetzt, dass der Wasserbehälter aufgrund des inneren Drucks einen Sprung bekommen hat und nicht wegen des Temperaturunterschiedes zwischen Hülle und Wasser. Auch in diesem Fall gilt, dass die Temperatur-Eigenschaft nur dann kausal wirksam am Zerbersten des Wasserbehältnisses beteiligt war,

²⁸⁵ Jackson, Frank & Pettit, Philip: „*Program Explanation – A General Perspective*“, in: *Analysis*, Vol. 50, 1990, S. 110.

wenn es auch die Impuls-Eigenschaft war – ergo trifft 3(i) zu. Die Temperatur des Wassers wiederum hat aber nicht in gleicher Weise den Impuls des Moleküls herbeigeführt wie dieser, wenn er kausal wirksam war, das Zerspringen des Behälters bewirkt hat. Das bedeutet aber nicht anderes, als dass auch in diesem Szenario 3(ii) gilt. Da auch keine Kombination aus Temperatur-Eigenschaft und Impuls-Eigenschaft gleichermaßen den Sprung im Gefäß hervorgerufen hat, ist auch in diesem Beispiel 3(iii) wahr. Die Kenntnis der relevanten Kausalgesetze und vollständiges Wissen über die Moleküle hätten ausgereicht, um das Zerspringen des Glasbehälters mit Sicherheit vorherzusagen.

In diesem Zusammenhang stellt sich die Frage, ob Annahme 3. plausibel ist. Für den Fall, dass F und G weder sequentielle noch gleichberechtigte Kausalfaktoren sind, ist es sicherlich einleuchtend, dass F und G nicht gleichermaßen kausal wirksam sein können. Wenn zudem präsupponiert wird, dass beide kausal wirksam sind, ist es sinnvoll anzunehmen, dass F nur in einem derivativen Sinn kausal wirksam sein kann. Die Beziehung zwischen der Instanziierung von F und dem Eintreten von *e* ist nachrangig gegenüber der Beziehung zwischen der Instanziierung von G und dem Eintreten von *e*. Denn ein Sich-Ereignen letzterer Beziehung hat mit Sicherheit das Eintreten der ersten Beziehung zur Folge – was umgekehrt jedoch nicht gilt. Ist es aber darüber hinaus sinnvoll anzunehmen, dass die F-Eigenschaft *in keinsten Weise* kausal wirksam am Herbeiführen von *e* beteiligt ist? Jackson und Pettit zufolge ist dies sowohl aus strategischen als auch aus theoretischen Gründen eine sinnvolle Annahme. Theoretisch betrachtet wäre es sinnlos F kausale Wirksamkeit zuzuschreiben, denn es ist unklar wie eine solche Form kausaler Wirksamkeit aussehen sollte. Schließlich würde die kausale Wirksamkeit, die eine Instanziierung von F angeblich ausübt, immer schon von der kausalen Wirksamkeit einer Instanziierung von G hinfällig gemacht werden. Wenn man zugesteht, dass eine Instanziierung von G *e* herbeigeführt hat, weil es eine Instanziierung von G war, dann kann eine Instanziierung von F schlichtweg nichts zum Eintreten von *e* beitragen, es sei denn F ist eine Ursache von G oder eine gleichberechtigte Ursache neben der Instanziierung von G. Selbst wenn die Instanziierung von F mit der Instanziierung von G identisch wäre, würde jene nicht in derselben Weise zum Eintreten von *e* beitragen wie diese. Schließlich würde die Instanziierung von F *e* nur verursachen, weil sie gleichzeitig auch eine Instanziierung von G ist. In anderen Worten das Haben von F würde nur deshalb einen kausalen Beitrag zum Sich-Ereignen von *e* leisten, weil es zugleich auch ein Haben von G ist. Zudem gibt es eine weitere theoretische Erwägung, aufgrund derer man der F-Eigenschaft in obigem Szenario durchaus jegliche kausale Wirksamkeit für das Eintreten von *e* absprechen kann. Genauer gesagt richtet sich dieses Bedenken generell gegen das Zuschreiben von kausaler Wirksamkeit an höherstu-

fige Eigenschaften. Schreibt man nämlich höherstufigen Eigenschaften kausale Wirksamkeit zu, scheint man nach Belieben andere kausal wirksame Eigenschaften erzeugen zu können. So wäre in einem solchen Fall etwa in Jackson und Pettits erstem Beispiel nicht nur die Zerbrechlichkeit des Glases kausal wirksam für das Zerbrechen des Glases gewesen, sondern auch die Eigenschaft, diese Eigenschaft zu haben – Meta-Zerbrechlichkeit, wenn man so will. Tatsächlich wäre dann auch das Haben dieser Eigenschaft (Meta-Zerbrechlichkeit) kausal wirksam für das Zerbrechen des Glases gewesen, also Meta-Meta-Zerbrechlichkeit. Es ist klar, dass sich auf diese Weise ad infinitum neue kausal wirksame Eigenschaften ersinnen ließen – was natürlich absurd ist. Ferner ist es auch aus strategischen Gesichtspunkten vernünftig 3. voll anzuerkennen, anstatt sowohl eine derivative als auch eine primitive Art kausaler Wirksamkeit vorauszusetzen. Denn in letzterem Fall müsste eine derivativ kausal wirksame Eigenschaft entweder kausal relevant sein oder nicht. Wenn sie nicht kausal relevant ist, wenn also nur primitiv kausal wirksame Eigenschaften einen kausalen Beitrag leisteten, dann scheint es als würden alle kausalen Kräfte auf die unterste Ebene, d.h. die Stufe der Elementarphysik, versickern.²⁸⁶ Damit wären alle Zuschreibungen von kausaler Relevanz in Annahmen und Aussagen der Spezialwissenschaften und auch des common sense hinfällig. Wenn derivativ kausal wirksame Eigenschaften hingegen kausale Relevanz besäßen, so müsste erklärt werden, wie sich diese zwei verschiedenen Arten kausaler Relevanz (einerseits die der primitiv kausal wirksamen und andererseits die der derivativ kausal wirksamen Eigenschaften) zueinander verhalten. Zudem müsste gezeigt werden, wieso verschiedene Erklärungen ein und desselben Ereignisses kausal wirksam auf unterschiedlichen Ebenen sind.²⁸⁷ Das eigentliche Problem entsteht für Jackson und Pettit jedoch erst, wenn man zu den drei bisher diskutierten Annahmen eine vierte hinzufügt:

4. The only way for a property to be causally relevant to the production of a certain effect is by being causally efficacious in the process of production.²⁸⁸

Es ist offensichtlich, dass Annahme 4. sozusagen eine Erweiterung von Hypothese 2. ist. Denn in 2. wurde behauptet, dass *eine* Möglichkeit für Eigenschaften kausal relevant zu sein darin besteht, kausal wirksam zu sein. Annahme 4. fügt dem hinzu, dass dies die *einzig*e Möglichkeit für Eigenschaften ist kausal relevant zu sein. Es ist offensichtlich, dass ein Zusammenwirken der vier oben genannten Annahmen große Auswirkungen auf die Möglichkeit

²⁸⁶ Vgl. Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003.

²⁸⁷ vgl. Jackson, Frank & Pettit, Philip: „Program Explanation – A General Perspective“, in: *Analysis*, Vol. 50, 1990, S. 110f.

²⁸⁸ Jackson, Frank & Pettit, Philip: „Program Explanation – A General Perspective“, in: *Analysis*, Vol. 50, 1990, S. 111.

kausaler Erklärungen hat. Jackson und Pettit fassen die Konsequenzen folgendermaßen zusammen:

This means that the four assumptions will drive us to dismiss the claims of the special sciences, and of course the claims of common sense, to be able to provide causal explanations in terms of causally relevant properties. For consider the sorts of properties invoked in those areas: the property of a group that it is cohesive; of a mental state that it is a belief that *p*; of a biological trait that it maximizes inclusive fitness. For each of these properties it is plausible that there is a property *G* lower down, so to speak, which is such that the higher up property *F* is efficacious only if *G* is, and yet the *F*-instance and the *G*-instance are neither sequential nor coordinate causal factors. But then by assumption 3, *F* is not causally efficacious, and by assumption 4 is not causally relevant. We shall have to regard all such properties as inefficacious, and so as irrelevant, and so as incapable of playing a role in causal explanation. The only properties with any claim to causal relevance and a proper place in causal explanations will be properties like mass and charge.²⁸⁹

Für den Fall, dass die 4. Annahme wahr ist, werden dadurch jegliche Ansprüche der Einzelwissenschaften und des common sense auf das Liefern von Kausalerklärungen mithilfe kausal relevanter Eigenschaften getilgt. Dies zeigt eine Analyse der Eigenschaften aus diesen Gebieten – wie etwa der Eigenschaft *F* eines mentalen Zustandes ein Glauben zu sein, dass *p* der Fall ist. Denn hier scheint es nämlich durchaus gerechtfertigt zu sein, dass es eine Eigenschaft *G* gibt, die auf einer niedrigeren (d.h. näher an der Elementarphysik gelegenen) Stufe als *F* steht und von der gilt, dass *F* nur dann kausal wirksam ist, wenn *G* kausal wirksam ist und zugleich *F* und *G* weder eine Kausalfolge noch gleichberechtigte Ursachen sind. Dann ist die Eigenschaft *F*, Annahme 3. zufolge, nicht kausal wirksam und nach 4. auch nicht kausal relevant. Die einzigen Eigenschaften, die einen legitimen Anspruch auf kausale Relevanz und eine Daseinsberechtigung in Kausalerklärungen zu haben scheinen, sind nunmehr physikalische (Grund-)Eigenschaften wie Masse oder Ladung.

Nach Jackson und Pettit sind aber auch diejenigen, die den Kausalerklärungen der Spezialwissenschaften und des common sense getrost den Rücken kehren, nicht gegen die gebündelten Auswirkungen der vier Annahmen gefeit. Dazu führen sie folgendes Beispiel ein:

Suppose that I explain the noise made by some mechanism by the property of the mechanism that some of its parts are loose. That property relates as *F* relates to *G* to the following more specific property: that this and that particular part are loose. It is the property, after all, of instantiating some such specific property, perhaps this, perhaps that, perhaps another one. Thus the explanation involving existential quantification – the reference to an indeterminate some – cannot be a proper explanation: it does not invoke an efficacious property and so does not invoke a property that is relevant to the noise. This lesson holds quite generally, so that many of the explanations which physicists would endorse must look suspect.²⁹⁰

Angenommen man wollte den Lärm, den eine Maschine verursacht, damit erklären, dass im Inneren der Maschine einige Teile locker sind, so stünde man vor einem ähnlichen Problem

²⁸⁹ Jackson, Frank & Pettit, Philip: „*Program Explanation – A General Perspective*“, in: *Analysis*, Vol. 50, 1990, S. 112.

²⁹⁰ Jackson, Frank & Pettit, Philip: „*Program Explanation – A General Perspective*“, in: *Analysis*, Vol. 50, 1990, S. 112.

wie oben beschrieben. Denn die Eigenschaft einige lockere Teile zu haben verhält sich wie F zu G zu der spezifischeren Eigenschaft, dass genau diese oder jene Teile locker sind. Genauer gesagt ist es die Eigenschaft eine solche spezifischere Eigenschaft zu instanziierten, die die Geräusche produziert (denn für den Lärm, den die Maschine verursacht spielt es zunächst keine Rolle, ob dieser durch die lockeren Teile A und B oder die losen Komponenten C und D verursacht wird). Da diese Art der Erklärung einen Existenzquantor enthält (dass es nämlich eine solche spezifischere Eigenschaft gibt), kann es keine angemessene Erklärung sein. Schließlich enthält sie keine kausal wirksame Eigenschaft und damit auch keine Eigenschaft, die kausal relevant für den Lärm wäre. Dies ließe sich auf die meisten Erklärungen in der Physik übertragen, wodurch selbst diese vermeintlichen Prototypen kausaler Erklärungen in einem neuen, dubiosen Licht erscheinen würden.

Auf die Frage, wie man mit diesen Ergebnissen umgehen sollte, antworten Jackson und Pettit wie folgt: „They would mean that if in any area we gain access to a lower-order explanation of something, then we should jettison higher-order explanations in its favour. That in itself seems unattractive.“²⁹¹ Nach Jackson und Pettit würde es also wenig Sinn machen, höherstufige Erklärungen immer zugunsten von niedrigstufigeren Erklärungen über Bord zu werfen. So könnte man zwar im Hinblick auf ihr erstes Beispiel durchaus gewillt sein die Zerbrechlichkeits-Erklärung auf die Frage nach dem Warum des Zersplitterns des Glases zurückzuweisen, wenn man Zugang zu der Erklärung durch die molekulare Struktur hat. Schließlich stünden jedem, der Einblick in Molekularstruktur-Erklärung hat, zugleich auch alle wichtigen Informationen zur Verfügung, die die Zerbrechlichkeits-Erklärung liefert. In den beiden anderen Beispielen scheinen die Dinge allerdings anders gelagert zu sein. Denn die Quadratische-Form-Erklärung und die Temperatur-Erklärung scheinen Informationen zu enthalten, die nicht schon allein aufgrund des Zugangs zu ihren niedrigstufigeren Gegenständen gewährleistet sind. Jemand, der zum Beispiel weiß, dass die Undurchdringlichkeit dieses Teils des Pfahls verhindert hat, dass dieser durch jene Öffnung geführt werden kann, weiß nicht notwendigerweise genausoviel wie jemand, der die Nicht-Einführbarkeit des Pflocks mit dessen quadratischer Grundform erklärt. Ganz ähnlich ist der Fall auch beim dritten Beispiel gelagert: jemand, der das Zerspringen des Glasbehälters mit dem Impuls dieses oder jenes Moleküls beim Auftreffen auf diese oder jene Molekülbindung der Hülle erklärt, vermittelt dadurch nicht unbedingt genauso viel an Information wie jemand, der das Zerbersten mit der steigenden Temperatur und dem damit einhergehenden anwachsenden Druck innerhalb des Behälters erklärt. Da man in vielen Fällen zudem keinen Zugang zu niedrigstufigen Erklärungen hat,

²⁹¹ Jackson, Frank & Pettit, Philip: „*Program Explanation – A General Perspective*“, in: *Analysis*, Vol. 50, 1990, S. 112.

müssten höherstufige Erklärungen in diesem Zusammenhang als Platzhalter für ‚echte‘ Kausalerklärungen betrachtet werden. Jemand, der also nur höherstufige Erklärungen kennt, hat somit *prinzipiell* keine Ahnung von den Eigenschaften, die eigentlich das erklären, womit er sich befasst. Denn die ihm bekannten Eigenschaften sind allesamt kausal irrelevant.²⁹²

Dies alles ist natürlich höchst unbefriedigend und daher eruieren Jackson und Pettit folgende Vorgehensweise: „The problem is to find (...) a ground for rejecting the only one of our four assumptions which looks questionable: assumption 4. Can we offer an account of causal relevance which allows an inefficacious property to be relevant to the production of an effect?“²⁹³

Da die Konjunktion der vier oben skizzierten Annahmen laut Jackson und Pettit zu der höchst verdrießlichen Situation führt, dass kausal nicht wirksame Eigenschaften gleichzeitig auch jeglicher kausaler Relevanz beraubt werden, müsse die einzige Annahme mit einem fragwürdigen Charakter, nämlich die vierte, zurückgewiesen werden. Zugleich müsse versucht werden ein Verständnis von kausaler Relevanz zu entwickeln, das es kausal nicht wirksamen Eigenschaften ermöglicht trotzdem kausal relevant zu sein. Genau darin besteht auch der von Jackson und Pettit vorgeschlagene Lösungsweg, den sie anhand des folgenden Beispiels illustrieren:

Consider the explanation why a piece of uranium emitted radiation over a certain period, which invokes the property of the uranium that some of its atoms were decaying: this, rather than the more specific property that such and such particular atoms were decaying. By our assumption 3, the property involving existential quantification cannot have been efficacious in producing the radiation. If it was efficacious, that is only because the more specific property was efficacious. And yet the instance of the abstract property did not relate to the instance of the more specific in the manner of a sequential factor or a coordinate one. So is there any other way in which the abstract property can have been causally relevant to the radiation, given that it was not causally efficacious?²⁹⁴

Stellt man sich die Frage, warum ein Stück Uran über einen gewissen Zeitraum radioaktive Strahlung abgegeben hat, so könnte man dies mit folgender höherstufigen Erklärung beantworten: der Klumpen Uran hat aufgrund der Eigenschaft des Urans, dass einige seiner Atomkerne zerfallen, gestrahlt. Dies wäre wohl auch eine bessere Antwort als (die niedrigstufige), dass genau diese und jene Atomkerne zerfallen sind – schließlich hätten auch andere Atomkerne zerfallen können und trotzdem hätte das Stück Uran radioaktive Strahlung emittiert. Nach Annahme 3. können aber Eigenschaften, die Existenzquantoren enthalten, nicht kausal wirksam sein. Wenn die Eigenschaft, dass einige Atomkerne zerfallen kausal wirksam war, dann nur deshalb, weil die spezifischere Eigenschaft, dass genau diese und jene Atomkerne

²⁹² vgl. Jackson, Frank & Pettit, Philip: „Program Explanation – A General Perspective“, in: *Analysis*, Vol. 50, 1990, S. 112f.

²⁹³ Jackson, Frank & Pettit, Philip: „Program Explanation – A General Perspective“, in: *Analysis*, Vol. 50, 1990, S. 113.

²⁹⁴ Jackson, Frank & Pettit, Philip: „Program Explanation – A General Perspective“, in: *Analysis*, Vol. 50, 1990, S. 114.

zerfallen kausal wirksam ist. Gibt es für die abstrakte Eigenschaft überhaupt eine Möglichkeit kausal relevant für die Strahlung zu sein? Jacksons und Pettits affirmative Antwort auf diese Frage lautet:

Yes, there is, and the answer is more or less obvious. Although not efficacious itself, the abstract property was such that its realization ensured that there was an efficacious property in the offing: the property, we may presume, involving such and such particular atoms. The realization of the higher-order property did not produce the radiation in the manner of the lower-order. But it meant that there would be a suitably efficacious property available, perhaps that involving such and such particular atoms, perhaps one involving others. And so the property was causally relevant to the radiation, under a perfectly ordinary sense of relevance, though it was not efficacious. It did not do any work in producing the radiation – it was perfectly inert – but it had the relevance of ensuring that there would be some property there to exercise the efficacy required.²⁹⁵

Es bestehe also durchaus die Möglichkeit, dass sich die abstrakte bzw. höherstufige Eigenschaft als kausal relevant erweist, obwohl sie selbst nicht kausal wirksam ist – und zwar, indem ihre Realisierung sicherstellt, dass eine kausal wirksame Eigenschaft eintritt. In oben genanntem Beispiel etwa könnte die Realisierung der abstrakten Eigenschaft, dass einige Atomkerne des Uranstücks zerfallen, die Instanziierung einer spezifischeren Eigenschaft, dass z.B. genau diese und jene Atomkerne zerfallen, garantieren. Es ist klar, dass die höherstufige Eigenschaft die Strahlung nicht in gleicher Weise hervorgebracht hat wie die niedrigstufige Eigenschaft. Allerdings bedeutet das Eintreten der höherstufigen Eigenschaft nichts anderes, als dass hiermit die Instanziierung einer geeigneten, niedrigstufigeren Eigenschaft, die kausal wirksam ist, gewährleistet wird. Dabei spielt es letztendlich keine Rolle welche spezifische Eigenschaft genau realisiert wird – also welche Atome genau zerfallen, ob es nun diese und jene oder völlig andere sind. Auf diese Weise ist die höherstufige Eigenschaft kausal relevant für die Strahlung des Uranklumpens, ohne dabei gleichzeitig auch kausal wirksam zu sein. Schließlich hat die abstrakte Eigenschaft überhaupt keine kausale Arbeit bei der Entstehung der Radioaktivität geleistet, aber sie war kausal relevant für die Emission der Strahlung in dem Sinn, dass sie die Instanziierung einer Eigenschaft sichergestellt hat, die kausal wirksam für das Freisetzen der Strahlung war.

Nun stellt sich die Frage, wie man sich die Beziehung zwischen einer solchen höherstufigen, kausal relevanten, wenngleich auch kausal nicht wirksamen, Eigenschaft und der Wirkung vorzustellen hat. Diese Beziehung wird von Jackson und Pettit folgendermaßen beschrieben:

The realization of the property ensures – it would have been enough to have made it suitably probable – that a crucial productive property is realized and, in the circumstances, that the event, under a certain description, occurs. The property-instance does not figure in the productive process leading to the event but it more or less ensures that a property-instance which is required for that process does figure. A useful metaphor for describing the role of the property is to say that its realization programs for the appearance of the productive property and, under a certain description, for the event produced. The analogy is with a computer program which ensures that

²⁹⁵ Jackson, Frank & Pettit, Philip: „*Program Explanation – A General Perspective*“, in: *Analysis*, Vol. 50, 1990, S. 114.

certain things will happen – things satisfying certain descriptions – though all the work of producing those things goes on at a lower, mechanical level.

The solution proposed for the problem we have been confronting is that in each case the higher-order inefficacious property is causally relevant to the event produced, because its realization programs for the realization of a lower-order efficacious property and, in the circumstances, for the occurrence of the event in question. The lower-order efficacious property may not be the lower-order property mentioned in each case but, if it is not, it will be one for which the realization of that property directly programs, one for which the realization of a property programmed for by that property directly programs, or whatever: it will be a property for which the original property programs indirectly, via the programming of intermediate properties.²⁹⁶

Die Realisierung der höherstufigen Eigenschaft stellt also sicher, dass eine niedrigstufigere, kausal wirksame Eigenschaft eintritt. ‚Sicherstellen‘ bedeutet in diesem Zusammenhang, dass die Implementierung der höherstufigen Eigenschaft allein schon genügt, um eine Verwirklichung der niedrigstufigeren Eigenschaft entsprechend wahrscheinlich zu machen. Obwohl die Instanziierung der höherstufigen Eigenschaft somit nicht am kausalen Hervorbringen des in Frage stehenden Ereignisses beteiligt ist, gewährleistet sie demnach mehr oder weniger, dass eine kausal produktive Eigenschaft realisiert wird. Eine Metapher, die die Rolle der höherstufigen Eigenschaft anschaulich charakterisiert, besagt, dass deren Realisierung die Instanziierung einer kausal wirksamen, niedrigstufigen Eigenschaft und somit auch das Eintreten eines bestimmten Ereignisses unter einer gewissen Beschreibung *programmiert*. Diese Analogie ist der Computerwelt entliehen – denn hier stellt die Implementation eines Programms sicher, dass gewisse Dinge passieren werden, die einer bestimmten Beschreibung gerecht werden. Allerdings liegen die kausalen Vorgänge, die die in Frage stehenden Dinge tatsächlich hervorbringen, auf einer anderen, d.h. niedrigeren und mechanischen, Ebene.

Ganz analog verläuft auch die von Jackson und Pettit vorgeschlagene Lösung für das von ihnen umrissene Problem. Denn höherstufige Eigenschaften können demzufolge, obwohl sie nicht *kausal wirksam* sind, dennoch *kausal relevant* für das Eintreten einer Wirkung sein. Dies ist genau deshalb der Fall, weil ihre Realisierung die Verwirklichung einer niedrigstufigeren, kausal wirksamen Eigenschaft programmiert. Die niedrigstufige Eigenschaft, die kausal wirksam ist, muss nicht zwangsläufig in jedem Fall dieselbe sein und kann sowohl direkt als auch indirekt (d.h. über intermediäre Eigenschaften) von der ursprünglichen Eigenschaft programmiert worden sein.

Um die Tauglichkeit ihres Lösungsvorschlages zu erproben, ist dessen Anwendung auf die drei obigen Beispiele angeraten, so Jackson und Pettit. Im Falle des ersten Beispiels stellt die Eigenschaft des Glases zerbrechlich zu sein sicher, dass das Glas eine bestimmte molekulare Struktur besitzt, die unabhängig davon wie diese Struktur genau realisiert ist, hinreichend dafür ist, dass das Glas unter bestimmten Umständen zerbricht. Im zweiten Beispiel garantiert

²⁹⁶ Jackson, Frank & Pettit, Philip: „*Program Explanation – A General Perspective*“, in: *Analysis*, Vol. 50, 1990, S. 114f.

die quadratische Grundform des (undurchlässigen) Pflocks, dass es ein undurchdringliches Teil des Pfahls gibt, das ein Durchführen durch die Öffnung unmöglich macht. In diesem Fall kann schon aus einfachen geometrischen Erwägungen heraus geschlossen werden, dass der Pflock nicht durch die Öffnung passen wird. Schließlich gewährleistet auch im dritten Beispiel die Temperatur des Wassers mehr oder weniger – genauer gesagt mit nahezu hundertprozentiger Wahrscheinlichkeit – dass ein passend positioniertes Molekül einen Impuls haben wird, der ausreicht um eine Molekularbindung in der Behälterhülle brechen zu lassen und damit auch für einen Sprung im Behältnis sorgen wird.²⁹⁷

Schlussendlich unterscheiden Jackson und Pettit zwei Modi kausaler Relevanz für Eigenschaften, entweder sind diese kausal relevant „(...) through being efficacious in the production of whatever is in question, or through programming for the presence of an efficacious property.“²⁹⁸ Einerseits kann also die kausale Wirksamkeit einer Eigenschaft beim Hervorbringen einer Wirkung die kausale Relevanz der Eigenschaft absichern. Andererseits können Eigenschaften auch kausale Relevanz erlangen, indem sie die Realisierung einer kausal wirksamen Eigenschaft programmieren. Durch diese Erkenntnis kann auch das von Jackson und Pettit dargestellte Problem als gelöst betrachtet werden. Schließlich hat sich dadurch die problematische Annahme 4. als falsch erwiesen – es existiert eben nicht nur *eine einzige* Möglichkeit für Eigenschaften kausal relevant zu sein, nämlich indem sie kausal wirksam sind. Vielmehr können sie auch auf einem *zweiten* Weg kausale Relevanz erlangen und zwar indem sie die Realisierung einer kausal wirksamen Eigenschaft programmieren. Die von Jackson und Pettit vertretene Auffassung kausaler Relevanz ist offensichtlich schwächer als ihr Konzept kausaler Wirksamkeit. Nichtsdestotrotz muss für den Fall, dass eine Eigenschaft kausal relevant für das Eintreten einer Wirkung ist, immer auch eine (für die Wirkung) kausal wirksame Eigenschaft vorhanden sein. Diese ontologische *conditio* ist jedoch nicht mit einem epistemischen Imperativ gleichzusetzen, dass es immer möglich sein muss die kausal wirksame Eigenschaft auch identifizieren zu können. Tatsächlich sind „(...) perhaps most of the explanations we are ever likely to offer (...) program explanations.“²⁹⁹ Da also in vielen Kausalerklärungen die tatsächlich kausal wirksamen Eigenschaften unbekannt sind, handelt es sich beim Gros der Erklärungen um Programmerklärungen. Schließlich ist zu erwarten, dass die potentiellen Kandidaten für kausal wirksame Eigenschaften nur auf dem Level der Teilchen-

²⁹⁷ vgl. Jackson, Frank & Pettit, Philip: „*Program Explanation – A General Perspective*“, in: *Analysis*, Vol. 50, 1990, S. 115.

²⁹⁸ Jackson, Frank & Pettit, Philip: „*Program Explanation – A General Perspective*“, in: *Analysis*, Vol. 50, 1990, S. 115.

²⁹⁹ Jackson, Frank & Pettit, Philip: „*Program Explanation – A General Perspective*“, in: *Analysis*, Vol. 50, 1990, S. 116.

physik anzutreffen sind. Genauer gesagt würde eine Kenntnis der kausal wirksamen Eigenschaften den Umgang mit ganz bestimmten Mikropartikeln erfordern – allerdings ist ein Zugang zu solchen speziellen Elementarteilchen (bislang) nicht vorstellbar.³⁰⁰

Schließlich gilt es festzuhalten, dass es nicht der alleinige Zweck von Programmerklärungen ist zu erklären wie kausal nicht wirksame Eigenschaften relevant für die Verursachung eines Ereignisses sein können. Darüber hinaus haben Programmerklärungen auch durchaus eine weitere Funktion, denn eine „(...) program explanation of an event *e* may provide information which the corresponding process explanation does not supply.“³⁰¹ Die Existenzberechtigung von Programmerklärungen leitet sich also nicht von der von Prozessklärungen ab, sondern ist deshalb gegeben, weil Programmerklärungen manchmal Informationen enthalten, die in den entsprechenden Prozessklärungen nicht enthalten sind. Somit können Programmerklärungen nicht – zumindest nicht in jedem Fall – vollständig durch Prozessklärungen ersetzt werden und verfügen infolgedessen auch über einen Selbstzweck. Eine abschließende Analyse des Unterschieds zwischen Programm- und Prozessklärungen liefern Jackson und Pettit in folgendem Abschnitt:

According to David Lewis, to explain something is to provide information on its causal history. Let us interpret the causal history as a process, involving such and such efficacious properties, that leads to the event or whatever in question. A program explanation provides a different sort of information from that which is supplied by the corresponding process account and therefore a sort of information which someone in possession of the process account may lack. The process story tells us about how the history actually went: say that such and such particular decaying atoms were responsible for the radiation. A program account tells us about how that history might have been. It gives modal information about the history, telling us for example that in any relevantly similar situation, as in the original situation itself, the fact that some atoms are decaying means that there will be a property realized – that involving the decay of such and such particular atoms – which is sufficient in the circumstances to produce radiation. In the actual world it was this, that and the other atom which decayed and led to the radiation but in possible worlds where their place is taken by other atoms, the radiation still occurs.³⁰²

Nach David Lewis' Anschauung bedeutet ‚etwas erklären‘ nichts anderes als Informationen zu dessen kausaler Geschichte zu liefern. Wird die kausale Geschichte zudem als ein Prozess verstanden, der diese und jene kausal wirksame Eigenschaften enthält, die dieses oder jenes Ereignis hervorbringen, so könnte man den Unterschied zwischen Programm- und Prozessklärungen wie folgt herausarbeiten: eine *Prozessklärung* schildert wie die kausale Geschichte *tatsächlich verlaufen* ist, während eine *Programmerklärung* beschreibt wie die Geschichte *hätte verlaufen können*. Bezogen auf das Beispiel mit dem Stück Uran bedeutet das – die Prozessklärung der Strahlung führt aus, welche Atome genau zerfallen sind, so dass die

³⁰⁰ vgl. Jackson, Frank & Pettit, Philip: „Program Explanation – A General Perspective“, in: *Analysis*, Vol. 50, 1990, S. 115f.

³⁰¹ Jackson, Frank & Pettit, Philip: „Program Explanation – A General Perspective“, in: *Analysis*, Vol. 50, 1990, S. 116.

³⁰² Jackson, Frank & Pettit, Philip: „Program Explanation – A General Perspective“, in: *Analysis*, Vol. 50, 1990, S. 117.

radioaktive Strahlung zustande kam. Eine Programmerklärung wiederum würde sich damit begnügen zu sagen, dass unter vergleichbaren Umständen der Zerfall einer kritischen Menge irgendwelcher Atomkerne des Uranklumpens hinreichend für die Emission von Strahlung wäre. Wenn nämlich irgendwelche Atome zerfallen, ist damit auch sichergestellt, dass diese und jene (oder vollkommen andere) Atome zerfallen und damit ist auch eine Instanziierung einer kausal wirksamen Eigenschaft garantiert. In der aktuellen Welt waren es eben genau diese und jene Atome, die zerfallen sind und damit die Strahlung verursacht haben; in anderen möglichen Welten wird deren Platz von anderen Atomen eingenommen, die Radioaktivität des Uranstücks bleibt davon jedoch unberührt.

Tropen-Monismus

Ein weiterer Lösungsvorschlag für das Exklusionsproblem findet sich in David Robbs 1997 im *Philosophical Quarterly* erschienenen *The Properties of Mental Causation*.³⁰³ In diesem Aufsatz stellt David Robb zunächst fest, dass es einerseits zwar keine allgemein akzeptierte Theorie über das Wesen kausaler Relata gebe, dass aber andererseits die These Eigenschaften spielen eine wichtige Rolle in Kausalbeziehungen, eine weit verbreitete Annahme sei. So verursachen Objekte oder Ereignisse dieses oder jenes *allein aufgrund* der Instanziierung von gewissen Eigenschaften – diese Eigenschaften werden gemeinhin als kausal relevant bezeichnet. Ein klassisches Beispiel hierfür ist dieses: Eine Sopranistin singt das Wort ‚Zerspringe!‘ in einer sehr hohen Tonlage und bringt damit ein Glas zum Zerbersten. Obwohl das Wort, das die Sängerin intonierte ‚zerspringen‘ bedeutet, war dies *keine* kausal relevante Eigenschaft für das Zerspringen des Glases – eine solche ist vielmehr in der Tonlage zu sehen. Allgemein betrachtet stellt sich die Frage was Eigenschaften überhaupt kausal relevant macht. Dies ist insbesondere dann von Interesse, wenn es um mentale Eigenschaften geht. Schließlich werden diese für gewöhnlich als kausal relevant für Verhalten eingestuft. Allerdings haben mentale Eigenschaften gewisse Eigenheiten, die die angenommene kausale Relevanz rätselhaft erscheinen lassen. Erstens werden mentale Eigenschaften oftmals als nicht-physikalisch erachtet. Wenn nun aber die Physik ein in sich geschlossenes System ist, dann ist es nicht ersicht-

³⁰³ Auch in einem späteren Essay – nämlich dem 2003 im *American Philosophical Quarterly* erschienen *Mental Properties* – nimmt David Robb gemeinsam mit John Heil Stellung zum Wesen mentaler Eigenschaften. Anders als in *The Properties of Mental Causation* wird in *Mental Properties* jedoch kein konkreter Lösungsvorschlag für das Problem mentaler Verursachung hinsichtlich der kausalen Relevanz mentaler Eigenschaften formuliert. In dem Artikel geht es vielmehr um allgemeine Hintergrundannahmen im Hinblick auf Eigenschaften, die zum Exklusionsproblem führen.

lich wie etwas Nicht-Physikalisches in diesem System einen (kausalen) Unterschied machen könnte. Zweitens haben gewisse mentale Eigenschaften – und zwar intentionale – einen relationalen Charakter. Dies steht im Widerspruch dazu, dass nur die intrinsischen Eigenschaften eines Dinges bzw. Ereignisses relevant dafür sind, was von diesem Ding bzw. Ereignis verursacht wird. Der nicht-physikalische und relationale Charakter von mentalen Eigenschaften scheint diese also kausal wirkungslos im Hinblick auf Verhalten werden zu lassen. Diese epiphänomenalistische Bedrohung kommt Robb zufolge von zwei Seiten – und zwar von der nicht-physikalischen Front zum Einen und von der relationalen Front zum Anderen. Wobei Robb in *The Properties of Mental Causation* allein auf erstere Bezug nimmt.³⁰⁴

Die aktuelle Diskussion hinsichtlich mentaler Verursachung hat ihren Ausgangspunkt in Donald Davidsons Theorie des Anomalen Monismus, so Robb. Dieser Ansatz beruhe im Wesentlichen auf folgenden drei Thesen:

The principle of causal interaction: some mental events cause, and are caused by, physical events

The principle of the nomological character of causality: whenever two events are causally related, they are subsumed by a strict law

Psycho-physical anomalism: there are no strict psycho-physical laws.³⁰⁵

Nach Davidson wird die kausale Interaktion zwischen Mentalem und Physikalischem dadurch gewährleistet, dass jedes mentale Einzelereignis mit einem physikalischen Ereignis identisch ist. Dies hat Davidsons Theorie jedoch den Vorwurf eingebracht, dass sie letzten Endes die Art von Wirksamkeit des Mentalen leugnet, die sie eigentlich aufrechterhalten wollte. Denn die mentalen Eigenschaften eines Ereignisses sind nach dieser Auffassung kausal irrelevant für die besagte Wechselwirkung. Schließlich treten mentale Ereignisse demnach nur qua Physikalischem und nicht qua Mentalem in Kausalrelationen ein. Es ist offensichtlich, dass diese Sichtweise einen Epiphänomenalismus nach sich zieht und in klarem Widerspruch zu folgendem Prinzip (zu) steh(en schein)t:

Relevance: mental properties are (sometimes) causally relevant to physical events.³⁰⁶

Nach Robb kommt die Bedrohung des Prinzips ‚Relevance‘ aber nicht unbedingt von Davidsons Anomalismus, sondern vielmehr von zwei grundlegenderen metaphysischen Prinzipien – und zwar:

³⁰⁴ vgl. Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 178f.

³⁰⁵ Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 179.

³⁰⁶ Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 179.

Distinctness: Mental properties are not physical properties.

(...)

Closure: every physical event has in its causal history only physical events and physical properties.³⁰⁷

‚Distinctness’ besagt lediglich, dass mentale und physikalische Eigenschaften *numerisch* unterschiedlich sind. Das schließt nicht aus, dass alle mentalen Partikulare, wie z.B. Ereignisse, im Grunde physikalisch sind oder in anderen Worten mit physikalischen Ereignissen identisch sind. Zudem ist es nach ‚Distinctness’ auch durchaus möglich, dass mentale Eigenschaften von physikalischen Eigenschaften realisiert werden. Letzten Endes schreibt dieses Prinzip lediglich vor, dass dies *zwei Arten* von Eigenschaften sind und diese nicht mit jenen identisch ist, unabhängig davon wie eng ihre Beziehung ansonsten auch sein mag. Nach ‚Closure’ können nicht-physikalische Ereignisse und Eigenschaften an keinem Punkt in die physikalische Welt kausal eindringen. Daraus folgt offensichtlich, dass keine nicht-physikalische Eigenschaft kausal relevant für das Hervorbringen eines physikalischen Ereignisses sein kann. Da aber Verhalten eine physikalische ‚Äußerung’ ist und mentale Eigenschaften aufgrund von ‚Distinctness’ nicht mit physikalischen Eigenschaften identisch sein können, muss ‚Relevance’ in diesem Kontext zurückgewiesen werden und wieder scheint der Epiphänomenalismus zu obsiegen.

Ein weiteres Prinzip, das in diesem Zusammenhang eine Rolle spielt ist nach Robb:

Exclusion: if a property F’s being instantiated is causally sufficient for an event, then no property distinct from F is causally relevant to that event.³⁰⁸

Genauer gesagt kann man ‚Closure’ aus ‚Exclusion’ und einer weiteren Prämisse folgern. Diese Prämisse besagt, dass es für jedes physikalische Ereignis eine physikalische Eigenschaft gibt, deren Instanziierung kausal hinreichend für dessen Eintreten ist.

Zunächst scheint es folglich so, als wäre es unmöglich die drei genannten Prinzipien ‚Relevance’, ‚Distinctness’ und ‚Closure’ gemeinsam kohärent zu vertreten. Eine beliebte Lösung für das Problem wird von Robb wie folgt portraitiert:

This popular solution looks to supervenience as a way of reconciling (to some extent) *Relevance*, *Distinctness* and *Closure*. Although supervenience can be a relation between many sorts of entities, in this context it is usually thought to be a relation between mental and physical *properties*:

Supervenience: necessarily, for every *x* and every mental property *M* of *x*, *x* has some physical property *P* such that necessarily whatever has *P* has *M*.

This says, roughly, that every mental property is co-instantiated with a physical property that necessitates it. Proponents of the supervenience solution, as I shall call it, think that this relation will secure the efficacy of

³⁰⁷ Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 182f.

³⁰⁸ Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 184.

mental properties. Their reasoning seems to be that even though supervenience is a relation weaker than identity, and so compatible with *Distinctness*, it is strong enough to secure *Relevance* without violating *Closure*. Properties that supervene on the physical are supposed to be, in some way, ‘nothing over and above’ their subvenient bases, and so such properties can do their causal work *within* the closed physical world.³⁰⁹

Die Supervenienz-Lösung geht, in groben Zügen, davon aus, dass gleichzeitig mit jeder mentalen Eigenschaft eine physikalische Eigenschaft instanziiert wird und dass jene von dieser notwendig gemacht wird. Auf diese Art und Weise soll die kausale Wirksamkeit von mentalen Eigenschaften abgesichert werden. Der Hintergedanke scheint zu sein, dass Supervenienzbeziehungen einerseits schwächer als Identitätsrelationen und somit mit ‚Distinctness‘ vereinbar sind, andererseits aber auch stark genug um ‚Relevance‘ zu gewährleisten ohne dabei ‚Closure‘ zu verletzen. Denn Eigenschaften, die auf Physikalischem supervenieren, gehen in gewisser Weise nicht über ihre subvenienten Basen hinaus und so können sie ihre kausale Arbeit auch innerhalb der geschlossenen, physikalischen Welt leisten. Der Supervenienz-Lösungsansatz weist nach Robb zwar in die richtige Richtung, allerdings ist er auch mit einigen Problemen behaftet. Erstens muss die Supervenienz-Lösung streng genommen ‚Closure‘ und somit auch ‚Exclusion‘ zurückweisen. Wenn nämlich eine mentale Eigenschaft in der kausalen Geschichte eines physikalischen Ereignisses vorkommt, dann wird dadurch ‚Closure‘ verletzt, unabhängig davon wie eng mentale und physikalische Eigenschaften miteinander verknüpft sind (solange sie nicht identisch sind, d.h. falls ‚Distinctness‘ gilt). Aber selbst wenn die Supervenienz-Lösung dieses Problem umschiffen könnte, stünde sie vor einem zweiten Problem: warum sollte ausgerechnet Supervenienz eine adäquate Erklärung für mentale Verursachung unter diesen Vorzeichen liefern? Für gewöhnlich ist die Antwort auf diese Frage, dass es eine Vielzahl von Eigenschaften (nicht nur mentale, sondern auch biologische, psychologische, etc.) gibt, die auf physikalischen Eigenschaften supervenieren und trotzdem oder gerade deshalb kausal relevant sind. Dies soll gemeinhin als ‚Beweis‘ dafür gelten, dass mentale Eigenschaften ihre kausale Relevanz von physikalischen Eigenschaften erben. Robb hält diese Schlussfolgerung jedoch nicht für gerechtfertigt, sondern bilanziert sogar im Umkehrschluss, dass auch eben jene Eigenschaften, die in der Regel als kausal relevant erachtet werden es in Wirklichkeit gar nicht zwangsläufig sein müssen. Denn ‚Closure‘ scheint die kausale Relevanz *aller* nicht-physikalischen Eigenschaften in Frage zu stellen. Schließlich stünde der Supervenienz-Ansatz – selbst wenn man über die beiden bereits diskutierten Mängel hinwegsieht – noch vor einem dritten Problem. Es ist nämlich nicht ersichtlich *wie* Supervenienz es bewerkstelligen sollte, dass mentale Eigenschaften kausale Relevanz erlangen. Wenn es also rätselhaft ist auf welche Art und Weise nicht-physikalische Eigen-

³⁰⁹ Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 184f.

schaften in einer (kausal geschlossenen) physikalischen Welt kausale Arbeit verrichten könnten, dann stellt sich die Frage wie Supervenienz dabei behilflich sein könnte diesen Schleier zu lüften.³¹⁰

Robb glaubt, dass keines dieser Probleme der Supervenienz-Lösung den Garau bereiten kann und betrachtet seine eigene Lösung des Problems kausaler Relevanz mentaler Eigenschaften – den Tropen-Monismus – auch als eine Spielart der Supervenienz-Lösung. Ihren Ausgang nimmt Robbs Lösung an einer Unterscheidung mehrerer Gebrauchweisen der Wortes ‚Eigenschaft‘ in diesem Zusammenhang. Dazu Robb:

There are two ways to read ‘property’. One is as ‘universal’ or, for the nominalist, ‘class’. On this reading, the property F is the universal F-ness or the class of all Fs. Properties in this sense are unifying entities. They are what all Fs have in common: either the Fs all share a universal or they all belong to the same class. (...) But the other reading of ‘property’ is as ‘abstract particular’ or, as I would prefer, ‘trope’. On this reading, properties are particulars, wholly present in the individuals that instantiate them but logically incapable of being (at the same time) wholly present elsewhere. An example may help to clarify this distinction. Two ripe bananas both have the property of being yellow. But is one banana’s colour numerically the same as the other’s? If you are thinking of properties in the first way, the answer is ‘Yes: they both share a universal or they are both members of the class of yellow things’. But if you are thinking of properties in the second way, the answer is clearly ‘No’. This banana’s yellowness is different from that one’s yellowness. Here we have two properties in different locations. These two colour tropes are of course similar, but they are distinct.³¹¹

Einerseits kann die Bedeutung von ‚Eigenschaft‘ also mit ‚Universalie‘ oder ‚Klasse‘ widergegeben werden. In diesem Sinn verkörpert beispielsweise die Eigenschaft F, die F-heit aller Partikulare, die die Eigenschaft F exemplifizieren oder in anderen Worten die Klasse aller Fs. Es ist offensichtlich, dass nach einem derartigen Verständnis ‚Eigenschaften‘ Entitäten mit einem stark vereinigenden Charakter sind. Die Eigenschaft F ist dann nichts anderes als das, was alle Objekte, die F instanzieren, gemeinsam haben. Andererseits ist es aber auch durchaus möglich unter ‚Eigenschaft‘ ein abstraktes Partikular oder – in Robbs Nomenklatur – ‚Trope‘ zu verstehen. In diesem Sinne sind Eigenschaften Partikulare, die sich zum Zeitpunkt der Instanzierung der Eigenschaft gänzlich in dem Objekt, das die Eigenschaft gerade instanziiert, befinden. Folglich ist es für eine Eigenschaft nach diesem Verständnis, ganz im Gegensatz zur ersten Lesart des Wortes, logisch unmöglich zum Zeitpunkt ihrer Instanzierung gleichzeitig in einem beliebigen anderen Objekt präsent zu sein. Zur Veranschaulichung dieser Unterscheidung führt Robb folgendes Beispiel an: zwei reife Bananen besitzen beide die Eigenschaft gelb zu sein. Auf die Frage, ob die Farbe der einen Banane numerisch identisch ist mit der Farbe der anderen Banane, gibt es laut Robb zwei Antworten. Versteht man nämlich ‚Eigenschaft‘ in dem ersten Sinn (also als Universalie oder als Klasse) lautet die Antwort

³¹⁰ vgl. Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 185f.

³¹¹ Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 186f.

offenkundig ‚Ja‘. Schließlich gehören beide Bananen der Klasse der gelben Dinge an, sie exemplifizieren dieselbe Universalie. Folgt man aber wiederum der zweiten Taxonomie von ‚Eigenschaft‘ stellt sich die Sachlage gänzlich anders dar. Denn in diesem Fall müsste auf die Frage ein eindeutiges ‚Nein‘ erwidert werden, weil es sich nämlich um zwei verschiedene Gelb-Tropen handelt. Schon allein durch die räumliche Distanz zwischen beiden Bananen können die beiden Eigenschaften gelb zu sein nach dieser Auslegung von ‚Eigenschaft‘ *nicht* identisch sein. Zwar sind die beiden Gelb-Tropen einander gewiss ähnlich, können aber per definitionem nicht identisch sein. Um die beiden Lesarten von ‚Eigenschaft‘ eindeutig differenzieren zu können, führt Robb folgende Begriffe zur Unterscheidung ein: „In what follows, it will be useful to have two different terms for these two quite different kinds of properties. I shall use the term ‘type’ to refer to universals or classes, but I shall use the term ‘trope’ to refer to abstract particulars. When a given trope is a trope of type F, I shall call it an F-trope. I shall reserve the term ‘concrete particular’ for the things (events, objects, etc.) that *have* tropes and are *of* types – so, for example, the two bananas are both concrete particulars that have their own yellow-tropes and are of the same type, namely yellow.“³¹²

Auf dieser Unterscheidung zwischen Tropen und Typen beruht auch Robbs Lösungsansatz für das Exklusionsproblem hinsichtlich mentaler Verursachung. Seine Strategie zur Bewältigung des Problems formuliert er wie folgt:

With tropes I think we can reconcile *Distinctness* and *Closure* with *Relevance*.

We can do this by first formulating a version of monism: I shall call it ‘trope monism’. Like Davidson’s view, trope monism says that all mental events are physical events even though mental and physical types are distinct. But it goes beyond Davidson in recognizing intermediate entities, mental tropes, and identifying them with physical tropes. Although this may seem like an insignificant addition, it is crucial to solving the present problem. Although tropes are not types (i.e. universals or classes), they fill a role traditionally assigned to types: they characterize particulars such as objects and events. And they also fill another role that types are often thought to play: they are the ‘properties’ that are causally relevant in causal relations; they are, as I shall say for short, the properties of causation. This allows us to piece together a solution to the problem. When *Relevance* says that mental ‘properties’ are causally relevant to physical events, what this means is that mental tropes are relevant. But trope monism says that mental tropes are physical, so we have not violated *Closure*. Nor have we violated *Distinctness*: mental and physical types are not the same even though every mental trope has a physical one, for trope monism does not rule out the multiple realizability of mental types.

(...)

This trope solution, as I shall call it, thus claims that the three principles are not inconsistent after all. Though ‘property’ should be read as type (i.e., universal or class) in *Distinctness*, it should be read as trope in *Relevance* and *Closure*. Here then are the principles once these two sorts of ‘properties’ are distinguished:

Relevance: mental tropes are (sometimes) causally relevant to physical events

Distinctness: mental types are not physical types

Closure: every physical event has in its causal history only physical events and physical tropes.

(*Exclusion*, one of the premises used in support of *Closure*, also gets a trope reading: if trope T’s being instantiated is causally sufficient for an event, then no trope distinct from T is causally relevant to that event.)

³¹² Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 187.

This, I claim, is the most plausible way to render these three principles consistent.³¹³

Für Robb besteht das Exklusionsproblem also im Kern darin die Prinzipien *Distinctness* und *Closure* mit *Relevance* in Einklang zu bringen. In seinem sogenannten Tropen-Monismus sieht er den Königsweg zu diesem Ziel. Monistisch ist Robbs Position deshalb, weil er wie Davidson davon ausgeht, dass alle mentalen Ereignisse letztendlich mit physikalischen Ereignissen identisch sind; mentale Typen sind jedoch nicht (unbedingt) identisch mit physikalischen Typen. Robbs Ansatz geht aber über Davidsons Theorie insofern hinaus, als dass er eine neue Art von Entitäten einführt – nämlich Tropen, wobei mentale Tropen identisch mit physikalischen Tropen sind. Zwar sind Tropen nicht identisch mit Typen, dennoch haben sie gewisse Funktionen, die für gewöhnlich Typen zugeschrieben werden. So kennzeichnen Tropen etwa Partikulare wie Objekte oder Ereignisse. In anderen Worten sie können bei der Identifikation von Partikularen hilfreich sein. Zudem sind Tropen die (Art von) ‚Eigenschaften‘, die in Verursachungsbeziehungen kausal relevant sind. Auf diese Weise wird auch Robbs Lösung für das Exklusionsproblem evident: wenn nämlich in *Relevance* mentale Eigenschaften als kausal relevant für physikalische Ereignisse dargestellt werden, so heißt das nichts anderes, als dass mentale Tropen kausal relevant sind. Da nun aber mentale Tropen definitionsgemäß mit physikalischen Tropen identisch sind, wird somit auch *Closure* gewahrt. Schließlich trifft auch *Distinctness* zu, denn mentale Typen sind von physikalischen Typen verschieden, obwohl jedes mentale Trope identisch mit einem physikalischen Trope ist. Nach Robb ist es mittels des Tropen-Monismus’ möglich die drei Prinzipien *Relevance*, *Closure* und *Distinctness* konsistent miteinander zu vertreten. Dazu ist nur nötig, explizit zu machen was im jeweiligen Kontext genau unter ‚Eigenschaft‘ zu verstehen ist. Während *Distinctness* die erste Lesart von ‚Eigenschaft‘, also als Universalie oder Klasse, zugrunde legt, sollte ‚Eigenschaft‘ in *Relevance* und *Closure* als Trope verstanden werden. Folglich besagt *Relevance* dann, dass mentale Tropen (manchmal) für physikalische Ereignisse kausal relevant sind. *Distinctness* hingegen konstatiert, dass mentale Typen nicht identisch mit physikalischen Typen sind. Und *Closure* drückt schließlich aus, dass jedes physikalische Ereignis in seiner kausalen Geschichte ausschließlich physikalische Ereignisse und physikalische Tropen hat.

Obwohl der Tropen-Monismus laut Robb sozusagen der Gattung der Supervenienztheorien zuzurechnen ist, ist er dennoch nicht mit den oben diskutierten Schwierigkeiten dieser Ansätze konfrontiert. So ist der Tropen-Monismus Robb zufolge durchaus kompatibel mit *Closure* – denn ein kausal relevantes mentales Trope muss immer mit einem physikalischen Trope

³¹³ Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 187f.

identisch sein. Zudem gelingt es mithilfe des Tropen-Monismus den mysteriösen Schleier, der Erklärungsversuche mentaler Verursachung bei Supervenienztheorien normalerweise umgibt, zu lüften. Da nämlich mentale Verursachung im Tropen-Monismus im Grunde ihrer Natur nach nichts qualitativ anderes ist als physikalische Verursachung, gelingt es dieser Position auch die Frage nach dem ‚wie‘ mentaler Verursachung, die Vertreter von Supervenienztheorien in diesem Zusammenhang für gewöhnlich schuldig bleiben, adäquat zu beantworten.³¹⁴

Schließlich diskutiert Robb in *The Properties of Mental Causation* auch noch zwei mögliche Einwände gegen den von ihm formulierten Tropen-Monismus. Die erste mögliche Entgegnung formuliert Robb so:

(1) First, one might object that the trope solution's focus on tropes, not types, leads to the same problem that plagued Davidson's view. The problem with Davidson's view was the even if mental events are identical with physical events, mental events do not cause behaviour *qua* mental, but only *qua* physical. Similarly, against the trope solution one might object that even if mental events do cause behaviour in virtue of their mental tropes, they do not do so in virtue of those tropes being mental, but only in virtue of their being physical. That is, such tropes are not relevant in virtue of being tropes of mental types.³¹⁵

Das erste Bedenken, das laut Robb gegen den Tropen-Monismus geäußert werden könnte, gleicht in seiner Grundstruktur einer weit verbreiteten Kritik an Davidsons Anomalem Monismus. Diese besagt kurzum, dass selbst wenn mentale Ereignisse identisch mit physikalischen Ereignissen sind, dadurch noch nicht die kausale Kraft von mentalen Ereignissen bewiesen ist. Schließlich sind die in Frage stehenden mentalen Ereignisse schlichtweg ausschließlich deshalb kausal relevant, weil sie gleichzeitig auch physikalische Ereignisse sind. Diese Ereignisse verursachen Verhalten also nicht in ihrer Eigenschaft als mentale Ereignisse oder in anderen Worten *qua* Mentalem, sondern sie tun dies allein in ihrer Eigenschaft als physikalische Ereignisse oder *qua* Physikalischem. Es ist offensichtlich, dass sich diese Form der Kritik auch sehr einfach auf den Tropen-Monismus übertragen lässt. Schließlich könnte analog erklärt werden: selbst wenn mentale Ereignisse Verhalten aufgrund eines mentalen Trope verursachen, so tun sie diese nicht etwa deshalb weil das besagte Trope ein mentales Trope ist, sondern allein deshalb weil es definitionsgemäß auch gleichzeitig ein physikalisches Trope ist. Ganz allgemein gilt dann: kausal relevante Tropen sind nicht deshalb kausal relevant für das Erzeugen von Verhalten, weil sie einem mentalen Typ angehören, sondern vielmehr weil sie zeitgleich auch einem physikalischen Typ zuzurechnen sind. Ein Trope ist wiederum nicht als mentales Trope oder *qua* Mentalem kausal relevant; stattdessen verursacht es nur als physikalischen Trope oder *qua* Physikalischem Verhalten. Robbs Lösung für das

³¹⁴ Vgl. Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 189.

³¹⁵ Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 190.

‚qua-Problem‘ gleicht Davidsons Antwort auf diese Problematik. Denn jemand der einen solchen Einwand äußere, „(...) seems to be making a kind of category mistake. Tropes are not causally relevant *qua* this or that, they are causally relevant (or not), period.“³¹⁶ Im Gegensatz zu Davidsons Theorie ist eine Applikation des qua-Problems auf den Tropen-Monismus laut Robb nicht gerechtfertigt. Denn wie etwa das Beispiel der Sopranistin und des Glases zeigt, sind einige Eigenschaften eines Ereignisses kausal relevant für das, was das besagte Ereignis verursacht, während andere es nicht sind. Daher ist die zum Beispiel die Frage, ob ein mentales Ereignis Verhalten *qua* Mentalem oder *qua* Physikalischem verursacht in gewisser Weise begründet. Eine kausal relevante Eigenschaft hingegen lasse sich nicht in mehrere Aspekte zergliedern, von denen sich einige für die kausale Relevanz der Eigenschaft verantwortlich zeichnen und andere nicht. Diese gelte sowohl für Eigenschaften als Typen als auch für Eigenschaften als Tropen, so Robb. Folgendes Beispiel verdeutliche dies: ein roter Ball wird fallengelassen, trifft auf einer Metallplatte auf und hinterlässt eine Delle in der Platte. Nun kann man zwar sinnvollerweise fragen, ob der Ball die Delle *qua* Masse und nicht *qua* Röte verursacht hat, die Frage jedoch, ob die Masse des Balls *qua* Masse verursacht hat, erscheint sinnlos. Zudem drohe ein unendlicher Regress, wenn man eine Anwendung des qua-Problems auf Eigenschaften zuließe, so Robb. Aus diesen Gründen sieht Robb den Tropen-Monismus nicht durch das qua-Problem bedroht.

Der zweite mögliche Einwand gegen den Tropen-Monismus, den Robb diskutiert, ist folgender:

(2) A second objection says that making tropes, not types, the properties of causation makes too many properties causally relevant. An example of Yablo’s illustrates this objection:

Imagine a glass which shatters if Ella sings at 70 decibels or more. Tonight, as it happens, she sang at 80dB, with predictable results. Although, it was relevant to the glass’s shattering that the volume was 80 dB, it contributed nothing that it was *under* 90 dB.

The trope theorist wants to say in this case that there is *one* volume-trope present in Ella’s note. That one trope is at once an over-70-dB trope, an 80-dB trope and an under-90-dB trope. Since being over 70 dB and being 80 dB are surely causally relevant to the glass’s breaking, then the trope solution entails that being under 90 dB is also causally relevant to this effect. But this, the objection goes, cannot be right: being under 90 dB is not a causally relevant property here. So it looks as if we need to reject the trope solution: the causally relevant properties should be types, not tropes. This problem, of course, generalizes. A particular trope is a trope of many types. A given scarlet-trope, for example, is also a red-trope, a colour-trope, a beauty-trope (perhaps), etc. It is also a trope of uncountably many *disjunctive* types (if there are such things). It is, for example, a red-or-tall-or-angry-trope. If one event causes another in virtue of this trope, we do not want to say that all of the properties I mentioned are causally relevant; so it must be types, not tropes, that are relevant.³¹⁷

³¹⁶ Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 191.

³¹⁷ Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 191f.

Der zweite Kritikpunkt am Tropen-Monismus zielt darauf ab, dass diese Theorie zu einer wahren Inflation von kausal relevanten Eigenschaften führe. Um dies zu verdeutlichen sei nochmals das Sopranistinnen-Beispiel bemüht: nehmen wir also an für das Zerspringen eines Glases genügt es, wenn eine Sopranistin mit 70 dB oder mehr singt. Wenn nun eine Sopranistin zum Beispiel mit 80 dB singt, dann wird das Glas zerbersten. Dabei, so ist man wohl geneigt anzunehmen, ist es kausal relevant für das Zerschlagen des Glases, dass die Lautstärke 80 dB war. Allerdings ist es nicht kausal relevant, dass die Lautstärke unter 90 dB war. Der Tropen-Monismus verpflichtet zu der Aussage, dass genau *ein* Lautstärke-Trope in der Stimme der Sopranistin präsent war. Dieses Lautstärke-Trope ist nun aber zugleich ein über-70dB-Trope, ein 80dB-Trope und ein unter-90dB-Trope. Wenn nun aber das über-70dB-Trope und das 80dB-Trope kausal relevant für das Zerspringen des Glases sind, so müsste auch das unter-90dB-Trope kausal relevant dafür sein. Dies kann aber nicht der Fall sein, weil etwa auch ein 60dB-Trope ein unter-90dB-Trope ist und ein 60dB-Trope kann nicht kausal relevant für das Zerspringen des Glases sein. Aus diesem Grund müsse der Tropen-Monismus insgesamt zurückgewiesen werden, so der Vorwurf. Dieser Einwand gegen den Tropen-Monismus wiegt schwerer als der erste, so Robb, denn er ist intuitiv sehr einleuchtend. Allerdings gilt es die Denkweise hinter diesem Vorbehalt herauszuarbeiten. Zu diesem Zweck stellt Robb folgende Frage: Wenn c e verursacht, weil es heiß ist, warum ist man dann geneigt zu sagen, dass (i) c e verursacht, weil es heiß ist (was gleichbedeutend damit ist, dass in c ein Hitze-Trope ist) wahr ist, wohingegen (ii) c verursacht e, weil es heiß-oder-blau ist (ein Hitze-oder-Bläue-Trope in c ist) falsch ist? Zunächst gesteht Robb ein, dass dem Tropen-Monismus zufolge beide Behauptungen wahr sein müssen. Die Intuition, dass (i) wahr und (ii) falsch ist, rührt daher, dass sich (i) und (ii) essentiell darin unterscheiden, was sie nach sich ziehen. So beinhaltet (i) etwa, dass das Haben eines Hitze-Tropes hinreichend für die Verursachung von e ist, wahr ist, während (ii) zur Folge hat, dass das Haben eines Hitze-oder-Bläue-Tropes hinreichend für e ist, falsch ist. Denn wäre c blau gewesen (und gleichzeitig nicht heiß), dann hätte es e nicht verursacht, obwohl weiterhin ein Hitze-oder-Bläue-Trope in c gewesen wäre (aber selbstverständlich nicht derselbe Hitze-oder-Bläue-Trope wie im Beispiel). Die Intuition, dass (i) und (ii) einen unterschiedlichen Wahrheitswert haben, ist darin begründet, dass (i) und (ii) Unterschiedliches nach sich ziehen, so Robb – dies sei dadurch aber nicht bewiesen.³¹⁸

Eine letzte, äußerst radikale Variante der Kritik am Tropen-Monismus ist es zu bestreiten, dass es so etwas wie Tropen überhaupt gibt. Dementsprechend könne das Problem mentaler Verursachung auch nicht von mentalen Tropen gelöst werden. Schließlich gebe es nur Entitä-

³¹⁸ Vgl. Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 192f.

ten wie Typen oder konkrete Einzeldinge. Um diesem Vorwurf eine befriedigende Antwort entgegenhalten zu können, wäre eine tiefgehende Beschäftigung mit einer der Kernschwierigkeiten der Philosophie – dem Problem der Universalien – vonnöten. Robb schlägt stattdessen als bedingte Verteidigung von Tropen folgende Strategie vor: „If it turns out that the best way to explain mental causation requires, as I have argued, the recognition of tropes, then this is a good reason for believing that they exist. Indeed, perhaps the *only* reason for including an item in one’s ontology is the metaphysical work it can do, and fitting the mind into the physical world seems as important a metaphysical task as any.“³¹⁹ Für den Fall, dass die beste Lösung für das Problem mentaler Verursachung (im Sinne des Exklusionsproblems) die Existenz von Tropen beinhaltet, so sei dies durchaus ein legitimer Grund deren Existenz anzuerkennen, so Robb. Schließlich rechtfertigt vielleicht allein die metaphysische Arbeit, die ein Ding leistet, dessen Aufnahme in die Ontologie.

Mentale Quausalität

Terence Horgan publizierte im Jahr 1989 in der Zeitschrift *Philosophical Perspectives* den Aufsatz *Mental Quausation*, in dem er gleich mit dem ersten Satz sozusagen die ‚Gretchenfrage‘³²⁰ für nicht-reduktive Physikalisten stellt: „Even if individual mental events and states are causally efficacious, are they efficacious *qua* mental?“³²¹ Sollte man also gewillt sein mentalen Ereignistoken oder Zuständen kausale Wirksamkeit zuzuschreiben, dann stellt sich trotzdem die Frage, ob besagte Ereignistoken oder Zustände *qua* Mentalem – also aufgrund ihrer Zugehörigkeit zu einem mentalen Typ bzw. einer mentalen Eigenschaft – kausal wirksam sind. Anschließend stellt Horgan seine Agenda zur Beantwortung dieser Fragestellung vor:

My project here has four stages. First, I shall argue that our common-sense belief in the efficacy of the mental presupposes a positive answer to the question just posed (...). Second, I shall consider and reject several possible accounts of what it is for two events *c* and *e*, together with two properties *F* and *G*, to jointly instantiate the 4-place relation expressed by the locution ‘*c qua F* causes *e qua G*’. (I call this relation *quausation*.) Third, I shall propose – albeit somewhat sketchily – a positive account of quausation. And fourth, I shall argue that this account, despite its sketchiness, makes very plausible an affirmative answer to the question of the causal efficacy of the mental *qua* mental.³²²

³¹⁹ Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997, S. 194.

³²⁰ Diese Frage bringt gleichermaßen die Quintessenz des Exklusionsproblems zum Ausdruck.

³²¹ Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 47.

³²² Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 47.

Zunächst will Horgan in seinem Artikel aufzeigen, dass man nach dem common-sense Verständnis auf eine affirmative Antwort hinsichtlich Frage nach der kausalen Wirksamkeit des Mentalen qua Mentalem festgelegt zu sein scheint. Schließlich sei es in der alltäglichen explanatorischen Praxis gang und gäbe dem Mentalen kausale Wirksamkeit zuzusprechen. Zweitens diskutiert Horgan einige Konzepte von ‚Quausalität‘, die er aber allesamt zurückweist.³²³ Wobei er mit dem Neologismus ‚Quausalität‘ eine vierstellige Relation zwischen zwei Ereignissen *c* und *e* und zwei Eigenschaften *F* und *G* bezeichnet, deren Beziehung zueinander mit dem Satz ‚*c* *qua* *F* verursacht *e* *qua* *G*‘ ausgedrückt werden kann. Als drittes präsentiert Horgan seinen persönlichen Ansatz zur Quausalität, der ihm viertens eine positive Antwort auf die Ausgangsfrage erlaube.

Ausgehend von Davidsons Handlungstheorie erläutert Horgan das Problem mentaler Quausalität. Nach Davidson könnte jemand, der einem bestimmten primären Grund für eine Handlung hat und diese Handlung auch vornimmt, die Handlung *nicht* wegen dieses Grundes vornehmen. Im folgenden Satz werde z.B. etwas benötigt, um das erste ‚und‘ in ein ‚weil‘ zu verwandeln.

- (1) Er trainierte und er wollte abnehmen und dachte, dass Training dies bewirken würde.

Nach Davidson genüge es den Satz folgendermaßen umzuwandeln:

- (2) Sein Training wurde auf eine angemessene Art und Weise dadurch verursacht, dass er ein Verlangen hatte abzunehmen und dem Glauben, dass das Training dies bewirken würde.

Kurzum behauptet Davidson also, dass die Wahrheit von (1) und (2) notwendig und hinreichend für die Wahrheit von

- (3) Er trainierte, weil er abnehmen wollte und glaubte, dass Training dies bewirken würde.

ist. Dem hält Horgan entgegen, dass (1) und (2) zwar offensichtlich notwendige, jedoch keine hinreichenden Bedingungen für die Wahrheit von (3) sind – wie etwa die folgenden beiden,

³²³ In der folgenden Zusammenfassung von Horgans Argument wird dessen Beschreibung und kritische Diskussion der unterschiedlichen Konzepte von Quausalität weitestgehend ausgeklammert, stattdessen wird das Hauptaugenmerk auf Horgans eigene Theorie von Quausalität gelegt.

einschlägig bekannten, Beispiele demonstrieren.³²⁴ Das erste Szenario ist das Sopranistinnen-Beispiel – ein Sängerin bringt mit dem Gesang des Wortes ‚Zerspringe‘ ein Glas zum Zerschlagen. Dazu Dretske:

Meaningful sounds, if they occur at the right pitch and amplitude, can shatter glass, but the fact that these sounds have a meaning is surely irrelevant to their having this effect. The glass would shatter if the sounds meant something completely different or if they meant nothing at all. This fact doesn't imply that the sounds don't have a meaning, but it *does* imply that their meaning doesn't help explain their effects on the glass. To know *why* the glass shattered, you have to know something about the amplitude and frequency of these sounds, properties of the sound that are relevantly involved in its effect on the glass.³²⁵

Während also die Bedeutung des gesungenen Wortes für das Zerschlagen des Glases keine (kausale) Rolle spielt, sind Tonlage und Schwingungsbreite kausal relevante Eigenschaften für das Zerschlagen des Glases, so Dretske. Was, bzw. ob das intonierte Wort überhaupt etwas bedeutet, ist für die Wirkung des Gesangs unwesentlich; so wäre das Glas auch dann zerbrochen, wenn der Gesang etwas völlig anderes oder auch gar nichts bedeutet hätte. Für das Wissen warum das Glas zerbrochen ist, ist die Bedeutung des intonierten Wortes nach Dretske folglich irrelevant.

Ein zweites, ähnliches Beispiel findet sich bei Ernest Sosa:

A gun goes off, a shot is fired, and it kills someone. The loud noise is the shot (...) [I]n a certain sense the victim is killed by the loud noise; not by the loud noise as a loud noise, however, but only by the loud noise as a shot, or the like (...) [T]he loudness of the shot has no causal relevance to the death of the victim: had the gun been equipped with a silencer the shot would have killed the victim just the same.³²⁶

In Sosas Gedankenexperiment wird eine Person erschossen. Die Tatwaffe gibt dabei einen lauten Knall ab. In gewisser Weise kann man nun behaupten, dass der laute Knall der Schuss ist und somit die Person getötet habe – aber eben nicht als lauter Knall, sondern nur als Schuss. Es ist offensichtlich, dass das Lautsein des Schusses keine kausale Relevanz für den Tod der Person hatte. Denn wäre die Waffe etwa mit einem Schalldämpfer ausgestattet gewesen, hätte es keinen lauten Knall gegeben, das Opfer wäre aber nichtsdestotrotz gestorben.

Diese beiden Szenarien sollen verdeutlichen, dass in Davidsons Beispiel die Wahrheit von (1) und (2) nicht hinreichend für die Wahrheit von (3) ist. Angenommen die Sätze (1) und (2) sind wahr. Zugleich sei aber die Eigenschaft der Ursache ein primärer Grund zu sein genauso irrelevant für die Beziehung zwischen Ursache und Wirkung wie es die Bedeutung des Gesangs hinsichtlich des Zerschlagens des Glases, bzw. die Lautstärke des Schusses angesichts des Todes der Person, ist. Dann nämlich ist (3) schlichtweg falsch.

³²⁴ Vgl. Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 48f.

³²⁵ Dretske, Fred: „*Reasons and Causes*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 1f.

³²⁶ Sosa, Ernest: „*Mind-Body Interaction and Supervenient Causation*“ in: *Midwest Studies in Philosophy*, Vol. 9, 1984, S. 277f.

Um das Problem der mentalen Quausation in Davidsons Beispiel in den Griff zu bekommen führt Horgan Bedingung (4), die – wenn sie wahr ist – sowohl notwendig als auch hinreichend für die Wahrheit von (3) ist.

- (4) Sein Trainieren, *qua* der Eigenschaft *ein Trainieren zu sein*, wurde verursacht von seinem Wunsch abzunehmen und seinem Glauben, dass Training dies bewirken würde, *qua* der jeweiligen Eigenschaften *ein Wunsch abzunehmen zu sein* und *ein Glaube, dass das Training ein Abnehmen bewirkt, zu sein*.³²⁷

Dementsprechend definiert Horgan die Quausalitätsrelation so:

For any two events *c* and *e* and any two properties *F* and *G*, *c qua F* causes *e qua G* iff:

- (i) *c* causes *e*;
- (ii) *c* instantiates *F*;
- (iii) *e* instantiates *G*; and
- (iv) the fact that *c* instantiates *F* is explanatorily relevant to the fact that *e* occurs and instantiates *G*.³²⁸

Die Bedingung (iv) bringt laut Horgan zum Ausdruck, warum die Wahrheit von Äußerungen wie (1) und (2) nicht für die Wahrheit des Satzes (3) hinreichend sein kann. Diese besondere Art der explanatorischen Relevanz bezeichnet Horgan als ‚quasale Relevanz‘.

Wird nun die Eingangsfrage nach dem quausalen Relevanz mentaler Eigenschaften negiert, so hätte dies katastrophale Auswirkungen auf das Alltagsverständnis von der kausalen Wirksamkeit des Mentalen. Das common-sense Verständnis der kausalen Wirksamkeit des Mentalen fasst Horgan so zusammen: „We believe not merely that reasons *are* causes, but also that people act the way they do *because* they have reasons which rationalize their actions. More generally, we believe not merely that mental events and states *are* causes, but also that they have the effects they do *because* they instantiate the specific mental properties they do.“³²⁹

Der Glaube an die quasale Wirksamkeit mentaler Ereignisse und die quasale Relevanz mentaler Eigenschaften³³⁰ ist nach Horgan also tief im Selbstverständnis und Weltbild des Menschen verwurzelt. So werden Gründe gemeinhin als Ursachen verstanden und liefern zugleich ein gültiges Erklärungsmodell für das Verhalten von Individuen. Zudem gelten mentale Ereignisse gerade deshalb als Ursachen, weil sie gewisse mentale Eigenschaften instanziiieren.

³²⁷ Vgl. Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 50.

³²⁸ Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 50.

³²⁹ Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 51.

³³⁰ In gewisser Weise scheint Horgan bereits implizit Jacksons und Pettits Unterscheidung zwischen ‚kausaler Wirksamkeit‘ und ‚kausaler Relevanz‘ vorwegzunehmen, wie z.B. an folgendem Zitat besonders deutlich wird: „In short, common sense holds not merely that mental events and states are causally efficacious, but also that their mental properties are explanatorily relevant to the causal transactions in which those events figure as causes.“ (Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 51.)

Bestreitet man dies, drohe eine neue Form des Epiphänomenalismus – und zwar ein quausaler Epiphänomenalismus. Während der ‚traditionelle‘ Epiphänomenalismusvorwurf durch den Token-Physikalismus außer Kraft gesetzt werden kann, bedarf es für das Zurückweisen des quausalen Epiphänomenalismus nach Horgan zweier Schritte – nämlich: „(...) first, giving a plausible account of quausal relevance; and second, using this account to argue that the mental *qua* mental is causally efficacious.“³³¹

Ausgehend von Dretskes Sopranistinnen-Beispiel und von Sosas Schuss-Beispiel arbeitet Horgan sein eigenes Modell der quausalen Relevanz von Eigenschaften heraus. Zunächst stellt Horgan dabei Folgendes fest: ein Grund warum weder die Bedeutung des gesungenen Wortes, noch die Lautstärke des Schusses kausal relevant für das Eintreten des jeweiligen Wirkungsereignisses sind, liege in der Tatsache, dass die Instanziierung dieser beiden Eigenschaften keine kontrafaktisch notwendige Bedingung für das Eintreten der entsprechenden Wirkungsereignisse ist. Hätte die Sopranistin ein anderes Wort angestimmt (oder auch eine bestimmte Lautfolge ohne Bedeutung von sich gegeben) bzw. hätte der Todesschütze einen Schalldämpfer verwendet, wären die Wirkungsereignisse trotzdem geschehen. Daher muss laut Horgan also gelten:

(...) a counterfactual of the following form must be nonvacuously true:

- (N) If there had not occurred an event with the property F, then there would not have occurred an event with property G.³³²

Im Falle der Beispiele von Dretske und Sosa könnten die entsprechenden kontrafaktischen Bedingungen in etwa so aussehen:

- (N1) Wenn sich kein Geräusch, das ‚zerbrechen‘ bedeutet, ereignet hätte, wäre das Glas nicht zerbrochen.
(N2) Wenn sich kein lauter Schuss ereignet hätte, wäre das Opfer nicht gestorben.

Wenn nun (N1) und (N2) nicht nv-wahr³³³ sind, dann sind die Eigenschaften die Bedeutung-von-zerbrechen-zu-haben und die Eigenschaft ein-lauter-Schuss-zu-sein quausal nicht rele-

³³¹ Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 51.

³³² Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 56.

³³³ In obigem Zitat von Horgan ist die Rede davon, dass die kontrafaktische Bedingung „nonvacuously true“ sein muss. In diesem Kontext ist ‚nonvacuously‘ ein terminus technicus, der von David Lewis wie folgt definiert wird: „Given two propositions *A* and *C*, we have their *counterfactual* $A \Box \rightarrow C$: the proposition that if *A* were true, then *C* would also be true. This operation $\Box \rightarrow$ is defined by a rule of truth, as follows, $A \Box \rightarrow C$ is true (at a world *w*) iff either (1) there are no possible *A*-worlds (in which case $A \Box \rightarrow C$ is *vacuous*), or (2) some *A*-world where *C* holds is closer (to *w*) than is any *A*-world where *C* does not hold. In other words, a counterfactual is nonvacuously true iff it takes less of a departure from actuality to make the consequent true along with the ante-

vant für die jeweiligen Wirkungsereignisse. Allerdings lassen sich auch Hintergründe konstruieren, vor denen (N1) und (N2) nv-wahr sind. Wenn etwa die Sopranistin im ersten Beispiel fest entschlossen gewesen wäre entweder das Wort ‚Zerbreche‘ zu singen oder ihre Stimme überhaupt nicht zu erheben. Oder, wenn derjenige, der im zweiten Beispiel den Schuss abfeuert, keine Möglichkeit hatte den mit dem Schuss einhergehenden Lärm auf irgendeine Weise zu dämpfen. Folglich ist klar, dass es sich bei (N) zwar um eine notwendige, jedoch nicht um eine hinreichende Bedingung für die quausale Relevanz der F-Eigenschaften handelt. Um (N) dementsprechend zu stärken schlägt Horgan zunächst folgende Modifikation der kontrafaktischen Bedingung vor:

(N') If c had not instantiated property F, then there would not have occurred an event with property G,

where c is the event which was the cause in the actual world (or the “counterpart” of that event).³³⁴

Allerdings erweist sich nach einer eingehenderen Analyse auch die stärkere Variante der kontrafaktischen Bedingung als nicht haltbar. Betrachtet man nämlich in Dretskes Beispiel etwa diejenigen Eigenschaften, die sicherlich quausale Relevanz besitzen – also Tonlage und Schwingungsbreite (im Folgenden R-Eigenschaften) – und setzt sie in (N') erhält man:

(N1') Wenn c nicht die R-Eigenschaften besessen hätte, dann wäre das Glas nicht zerbrochen.

Nun stellt sich jedoch die Frage, ob (N1') tatsächlich nv-wahr ist (was der Fall sein müsste, wenn (N') nv-wahr ist)? Dies kann laut Horgan jedoch mit einer gewissen Berechtigung angezweifelt werden: schließlich könnte man behaupten, dass falls die Geräusche in einer anderen möglichen Welt nicht die R-Eigenschaften besäßen, diese Geräusche nicht mit den Lauten, die sich in der aktuellen Welt ereignet haben, identisch sind bzw. nicht legitimerweise als deren Pendant in der möglichen Welt gehandelt werden dürfen. Daher kommt Horgan auch zu dem Schluss, dass es ein Fehler ist die quausale Relevanz von Eigenschaften allein über kontrafaktische Bedingungen zu verstehen, deren Antezedens nur das Ursache-Ereignis der aktuellen Welt bzw. dessen Gegenstück in der möglichen Welt enthält. Vielmehr ist ein Ansatz

cedent than it does to make the antecedent true without the consequent.“ (Lewis, David: „*Causation*“ in: *Journal of Philosophy*, Vol. 70 (17), 1973, S. 560.) Da der Begriff ‚nonvacuously true‘ schwer ins Deutsche zu übersetzen ist, wird er von mir im Folgenden schlichtweg als ‚nv-wahr‘ wiedergegeben. Ein kontrafaktisches Konditional ist also genau dann nv-wahr, wenn es eine kleinere Abweichung von der aktuellen Welt darstellt das Antezedens und das Konsequenz gemeinsam wahr zu machen als das Antezedens wahr und das Konsequenz falsch zu machen.

³³⁴Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 57.

zu bevorzugen, der die prekäre Frage nach der Identität von Ereignissen über die mögliche Welten hinweg, außen vor lässt und sich stattdessen mit der Frage beschäftigt, was unter hinreichend ähnlichen Bedingungen passiert wäre, wenn ein Ereignis eingetreten (oder nicht eingetreten) wäre, das sowohl (a) die Eigenschaft F instanziierte als auch gleichzeitig (b) dem Ursache-Ereignis c in der aktuellen Welt angemessen ähnlich war.³³⁵

Für seine Erklärung der quausalen Relevanz von Eigenschaften referiert Horgan auf eine bestimmte Klasse möglicher Welten – nämlich die PSWs (*pertinently similar worlds*), das sind der aktuellen Welt angemessen ähnliche (mögliche) Welten:

Suppose that an event c causes an event e. Associated with the causal transaction, I suggest, is a set P[c,e] of *pertinently similar worlds* (PSW's). Each PSW contains a situation pertinent similar to – although perhaps somewhat different from – the situation in which c caused e in the actual world W. Also, each PSW w contains exactly one event c* which is *pertinently similar* to the event c of the actual world W; and w might, or might not, contain an event e* which is pertinent similar to the event e of W. An event in w, in order to be pertinent similar to c (or e) of W, not only must resemble c (or e) of W in pertinent intrinsic respects; it also must occur in *circumstances* in w what are pertinent similar to the circumstances in W in which the causal transaction between c and e transpired. If you like, the transworld pertinent-similarity relation holds between c*(of a PSW w) and c (of W) if (i) c* is a good *prima facie* candidate for being identical with (or the counterpart of) c; and (ii) c* is also a better such candidate than any other event in w. Likewise, *mutatis mutandis*, for the pertinent similarity between an event e* (of w) and e (of W). But, for the reason given above, it is best to allow the transworldly pertinent-similarity relation to be weaker than the transworld identity (or counterpart) relation.³³⁶

Angenommen ein Ereignis c verursacht ein Ereignis e, dann kann man diesem kausalen Vorgang laut Horgan eine Menge von möglichen Welten mit der Struktur P [c,e] zuordnen, die der aktuellen Welt angemessen ähnlich sind (im Folgenden werden diese Welten mit dem Kürzel PSW bezeichnet). Wobei jede PSW eine Situation enthält, die derjenigen in der aktuellen Welt, in der c e verursacht, angemessen ähnlich ist. Allerdings muss diese Situation in der PSW nicht exakt mit dem Vorgang in der aktuellen Welt, bei dem c e verursacht, identisch sein – es können also durchaus Unterschiede zwischen beiden Szenarien bestehen. Zudem enthält jede beliebige PSW w genau ein Ereignis c*, das dem Ereignis c in der aktuellen Welt W angemessen ähnlich ist. Schließlich kann es in w auch ein Ereignis e* geben, das dem Ereignis e in W angemessen ähnlich ist – dies muss jedoch nicht der Fall sein. Was bedeutet es nun für ein Ereignis in einer beliebigen PSW w dem Ereignis c (oder e) in der aktuellen Welt W angemessen ähnlich zu sein? Dazu muss ein Ereignis in w nicht nur hinsichtlich seiner intrinsischen Merkmale dem Ereignis c (oder e) in W angemessen ähneln, sondern es muss sich darüber hinaus auch unter Rahmenbedingungen ereignen, die denen von W angemessen ähnlich sind, unter denen die kausale Abfolge ‚c verursacht e‘ eingetreten ist. Folglich besteht zwischen einem Ereignis c* in einer PSW w und dem Ereignis c in der aktuellen Welt W eine

³³⁵ vgl. Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 57f.

³³⁶ Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 58.

angemessene Ähnlichkeit über mögliche Welten hinweg, wenn folgende beiden Kriterien erfüllt sind: Erstens muss c^* ein guter prima facie Kandidat für die Identität mit c sein, oder dafür, dessen Gegenstück in der möglichen Welt w zu sein. Und zweitens darf es in w kein anderes Ereignis geben, das die Rolle des prima facie Kandidaten besser ausfüllt als c^* . Da eine Identifizierung von Ereignissen über möglichen Welten hinweg generell eine diffizile Angelegenheit darstellt, ist man gut beraten die Beziehung der angemessenen Ähnlichkeit als eine, im Vergleich zur Identitäts- und Pendant-in-möglicher-Welt-Beziehung, schwächere Relation darzustellen, so Horgan.

Die genaue Definition von quausaler Relevanz im Rahmen von Horgans Konzeption der angemessenen ähnlichen Welten sieht folgendermaßen aus:

Quausal relevance of the properties of a cause-event and effect-event (...) depends upon what happens across the range of PSW's in the class $P[c,e]$. My proposal is this:

- If
- (i) event c causes event e ,
 - (ii) c and e respectively instantiate properties F and G ,
 - (iii) F and G are logically and metaphysically independent, and
 - (iv) the causal transaction between c and e does not involve preemption, overdetermination, or the like,
- then the fact that c and e instantiate F and G , respectively, is explanatorily relevant to the fact that c causes e iff the following *Relevance Condition* is satisfied:
- (R) For any world w in $P[c,e]$, if c^* is the event in w that is pertinently similar to c of W , then
 - (i) if c^* instantiates F in w , then c^* causes (in w) an event e^* which both instantiates G (in w) and is pertinently similar to the W -event e ; and
 - (ii) if c^* does not instantiate F in w , then c^* does not cause (in w) an event which is pertinently similar to the W -event e .³³⁷

Bei einer Verursachungsrelation zwischen zwei Ereignissen c und e , bemisst sich die quausale Relevanz der Eigenschaften F (von c) und G (von e) nach Horgan also daran, was über die Spannbreite der PSWs, die der Klasse $P[c,e]$ angehören, geschieht. Vorab müssen folgende vier Bedingungen erfüllt sein: erstens verursacht das Ereignis c das Ereignis e . Zweitens instantiiert das Ereignis c die Eigenschaft F und das Ereignis e instantiiert die Eigenschaft G . Drittens sind die Eigenschaften F und G logisch und metaphysisch voneinander unabhängig. Schließlich darf viertens die Kausalbeziehung zwischen c und e keine Präemption, Überdeterminierung, o.ä. beinhalten. Sind diese Grundvoraussetzungen erfüllt, dann ist die Tatsache, dass c F und e G instantiiert, explanatorisch relevant für die Tatsache, dass c e verursacht, wenn zugleich noch folgende Relevanzbedingung zutrifft. Die Relevanzbedingung formuliert Horgan anhand seiner Konzeption der angemessenen ähnlichen möglichen Welten mittels folgender Fallunterscheidung. Für jede mögliche Welt w , die der Klasse $P[c,e]$ angehört, gilt, wenn c^* das Ereignis in w ist, das dem Ereignis c in W angemessen ähnlich ist: erster Fall –

³³⁷Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 58f.

c^* instanziiert F in w . Dann verursacht c^* in w ein Ereignis e^* , das sowohl G in w instanziiert, als auch dem Ereignis e in W angemessen ähnlich ist. Zweiter Fall – c^* instanziiert F in w nicht. Dann verursacht c^* in w auch kein Ereignis, das dem Ereignis e in W angemessen ähnlich ist. In anderen Worten die Eigenschaften F und G der Ereignisse c und e sind genau dann quausal relevant für die Verursachung von e durch c , wenn für alle PSWs w gilt: entweder instanziiert ein Ereignis c^* in w (das dem Ereignis c in W angemessen ähnlich ist) die Eigenschaft F , dann verursacht c^* in w gleichzeitig ein Ereignis e^* , das G instanziiert und das dem Ereignis e in W angemessen ähnlich ist. Oder aber, c^* instanziiert nicht die Eigenschaft F , dann verursacht c^* allerdings auch kein Ereignis, das e in W angemessen ähnlich ist.

Nun stellt sich natürlich die Frage welchen Grad an Ähnlichkeit zur aktuellen Welt eine mögliche Welt (bzw. ein Ereignis darin) haben muss, um sich als PSW zu qualifizieren. Für eine adäquate Antwort darauf bedürfte es eigentlich eines besser ausgearbeiteten Ansatzes von quausaler Relevanz, so Horgan. Andererseits gebe es aber auch starke vor-theoretische Vorstellungen hinsichtlich der Frage der angemessenen Ähnlichkeit zwischen einer möglichen Welt und der aktuellen Welt, respektive der angemessenen Ähnlichkeit zwischen Ereignissen aus diesen beiden Welten. Diese intuitiven Urteile besitzen laut Horgan zunächst große Glaubwürdigkeit – sie könnten zwar jederzeit durch eine elaboriertere Theorie quausaler Relevanz widerlegt werden, solange dies jedoch nicht der Fall ist, sollte man seiner Eingebung vertrauen, so Horgan. Allerdings gibt es auch bestimmte mögliche Welten, bei denen man zunächst vielleicht geneigt wäre sie als PSW einzustufen, die Horgan jedoch ausdrücklich aus der Klasse der PSWs ausschließt. Hierbei handelt es sich um der aktuellen Welt sehr nahe mögliche Welten, in denen das Ursache-Ereignis c nicht die Eigenschaft F , sondern eine, der Eigenschaft F angemessen ähnliche, Eigenschaft F^* instanziiert, von der gilt: erstens gehört F^* derselben allgemeinen Art an wie F . Und zweitens ist F^* eine Eigenschaft, die wenn F quausal relevant wäre, im Wesentlichen auf dieselbe Art und Weise quausal relevant ist wie F .³³⁸

Eine konkrete Anwendung von seinem Konzept der quausalen Relevanz auf die Beispiele Dretskes und Sosas umschreibt Horgan folgendermaßen:

I am ready now to draw out some morals from the account of quausal relevance here suggested. To begin with, the account explains quite naturally why, in Dretske's example, the content of the sounds is irrelevant to their causing the glass to shatter. For, among the PSW's are worlds in which pertinently similar sounds occur but either mean something different or mean nothing at all; and the glass's shattering or not shattering is not correlated, across the PSW's, with the presence or absence of sounds which mean 'shatter'. (...)

Sosa's example of the shot's loudness fares similarly. Among the PSW's, relative to the shot's causing the victim's death, are worlds where although (i) there occurs a shot that is pertinently similar to the actual-world shot, nevertheless (ii) this shot is not loud – for instance, worlds in which a quieter gun is used, worlds where the

³³⁸ vgl. Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 60.

actual gun is fitted with a silencer, and so forth. The loudness of the actual shot is irrelevant, vis-à-vis the causal transaction in question, because of the presence or absence of a loud shot is not correlated across the PSW's, with the victim's dying or not dying.³³⁹

Laut Horgan illustriere seine Theorie der quausalen Relevanz von Eigenschaften auf geradezu paradigmatische Art und Weise, warum die Bedeutung des gesungenen Wortes bzw. die Lautstärke des abgefeuerten Schusses in den Beispielen Dretskes und Sosas nicht quausal relevant für das Eintreten der jeweiligen Wirkungsereignisse sein kann. Schließlich korreliere weder die Bedeutung der Worte noch die Lautstärke des Schusses über alle PSWs hinweg mit dem Zerschlagen des Glases, respektive dem Tod des Opfers. Denn es gibt einige PSWs, in denen die Sopranistin nicht ‚Zerschlage‘ singt, sondern ein anderes Wort intoniert oder auch Laute ohne jegliche Bedeutung von sich gibt, und in denen das Glas nichtsdestotrotz zerbricht. Auch hinsichtlich des Schusses gibt es PSWs, in denen kein lauter Knall mit dem Schuss einhergeht, z.B. weil eine leisere Waffe als in der aktuellen Welt verwendet wurde oder weil dieselbe Waffe wie in der aktuellen Welt mit einem Schalldämpfer ausgestattet war – auch in diesen PSWs wird das Opfer durch den Schuss getötet, es gibt dabei jedoch keinen lauten Knall. Kurzum, der ‚Lackmus-Test‘ für quausale Relevanz besteht laut Horgan also einfach darin, herauszufinden welche Eigenschaften über alle PSWs hinweg das jeweilige Wirkungsereignis begleiten – genau diese Eigenschaften sind quausal relevant für das Eintreten des Wirkungsereignisses, alle anderen sind zufällige Nebenprodukte (wie zum Beispiel die Bedeutung der gesungenen Worte oder die Lautstärke des Schusses), die mit den tatsächlichen Kausalvorgängen de facto nichts zu tun haben.

Im Unterschied zu den Beispielen Dretskes und Sosas ist bei typischen Fällen mentaler Verursachung die Sachlage etwas anders, so Horgan. Genauer gesagt ist hier ein Prinzip am Werk, das Horgan das ‚Prinzip der supervenienten Relevanz‘ nennt und wie folgt beschreibt:

Typically, when a mental event *c* causes another event *e*, the following principle of *Supervenient Relevance* holds for some salient mental property *F* of *c* and some salient property *G* of *e* that is logically and metaphysically independent of *F*:

(S.R.) For any world *w* in *P*[*c*,*e*], if *c** is the event in *w* that is pertinently similar to *c* of the actual world *W*, then (i) *c** instantiates *F* in *w*, and (ii) *c** causes (in *w*) an event *e** that both instantiates *G* and is pertinently similar to *e* of *W*.³⁴⁰

Nach dem Prinzip der supervenienten Relevanz gilt also für jede mögliche Welt *w*, die der Klasse *P* [*c*,*e*] angehört: wenn *c** das Ereignis in *w* ist, das dem Ereignis *c* in der aktuellen Welt *W* angemessen ähnlich ist, dann instanziiert *c** *F* in *w* und zugleich verursacht *c** in *w* ein Ereignis *e**, das sowohl *G* instanziiert, als auch dem Ereignis *e* in *W* angemessen ähnlich

³³⁹ Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 61.

³⁴⁰ Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 61.

ist. Es ist offensichtlich, dass das Prinzip der supervenienten Relevanz eine stärkere Form der Abhängigkeit bzw. Supervenienz zwischen den Eigenschaften F und G von verschiedenen c*- und e*-Ereignissen in unterschiedlichen PSWs postuliert als die einfache Relevanzbedingung in Horgans erster Definition quausaler Relevanz. So instanziiert ein c*-Ereignis in *jeder* PSW die Eigenschaft F und verursacht gleichzeitig ein e*-Ereignis, das die Eigenschaft G instanziiert. Zur Illustration gibt Horgan folgendes Beispiel an:

Consider, for example, a case where a mental event causes an action: Fred experiences a desire for a beer, and this desire causes him to walk to the refrigerator. The mental cause c has the property *being a desire for a beer*. Now, there are PSW's in which Fred experiences a desire for a beer in a situation that is modestly different from the actual situation – for instance, worlds in which Fred is located at a slightly different position in his house than his actual-world position at the time he experiences the desire for a beer. In these PSW's, Fred's desire causes a sequence of bodily motions that is somewhat different from his actual-world motions; however, the otherworldly sequence still qualifies as an action of walking to the refrigerator, and hence counts as pertinently similar to Fred's actual-world action. But are there any PSW's in which an event occurs which is both (i) pertinently similar to the actual-world event of Fred's desiring a beer, and yet (ii) lacks the mental property *being a desire for a beer*? I submit that there are not.³⁴¹

Dies ist sicherlich ein Paradebeispiel für einen Fall mentaler Verursachung: ein mentales Ereignis (Fred's Verlangen nach einem Bier) ist die Ursache für ein physikalisches Ereignis (Fred's Körperbewegungen, die ihn zum Kühlschrank führen). Nun gibt es sicherlich PSWs, in denen Fred das Verlangen nach einem Bier in Situationen verspürt, die sich leicht von den Umständen in der aktuellen Welt unterscheiden. So könnte Fred beispielsweise die Lust nach einem Bier an einem anderen Ort in seinem Haus (als in der aktuellen Welt) verspüren. In diesen PSWs würde Fred's Verlangen Körperbewegungen verursachen, die von denen in der aktuellen Welt divergieren. Trotzdem wären diese Körperbewegungen auch ein Gang zum Kühlschrank und daher angemessen ähnlich zu Fred's Handlung in der aktuellen Welt. Hinsichtlich des Ursacheereignisses könnte man sich wiederum fragen, ob es eine PSW gibt, in der ein Ereignis eintritt, das dem Ereignis der aktuellen Welt von Fred's Verlangen nach einem Bier angemessen ähnlich ist, das aber gleichzeitig nicht die mentale Eigenschaft ein Verlangen nach einem Bier zu sein besitzt. Horgan glaubt nicht, dass dies möglich ist. Zwar ist er durchaus geneigt mögliche Welten zuzulassen, in denen im Kopf von Fred's Doppelgänger ein Ereignis c* stattfindet, das physikalisch identisch ist mit dem Ereignis, das in der aktuellen Welt ein Verlangen nach einem Bier ist und das Körperbewegungen verursacht, die dem Gang zum Kühlschrank in der aktuellen Welt exakt gleichen, das jedoch *nicht* die Eigenschaft ein Verlangen nach einem Bier zu sein instanziiert. Allerdings unterscheiden sich solche mögliche Welten so fundamental von der aktuellen Welt, dass es unmöglich ist sie als PSWs zu betrachten, so Horgan. Um diese These zu untermauern, präsentiert Horgan zwei Strategien, die es

³⁴¹ Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 62.

gestatten sich eine mögliche Welt vorzustellen, die sich in diesem Umfang von der aktuellen Welt unterscheidet – nämlich eine *kopf-externe* und eine *kopf-interne* Variante. Bei der *kopf-externen* Version werden also alle Vorgänge in Freds Kopf und Körper bzw. dem Kopf und Körper von Freds Doppelgänger genauso belassen wie sie in der aktuellen Welt sind. Allerdings findet sich Fred gleichzeitig in einer Umwelt wieder, die sich in dem Maße von der aktuellen Welt unterscheidet, dass dort ein physikalisches Ereignis, das in der aktuellen Welt ein Verlangen nach Bier gewesen wäre, nicht mehr ein Verlangen nach einem Bier repräsentiert. Freds Doppelgänger könnte etwa ein Verlangen nach ‚Zwier‘ verspüren, obwohl er selbst diese Flüssigkeit als ‚Bier‘ bezeichnen würde. Zwier unterscheidet sich allerdings von Bier insoweit, dass es nicht aus Wasser, Hopfen und Malz, sondern aus einer chemischen Substanz mit der Grundstruktur XYZ³⁴², die sich äußerlich und in ihren intrinsischen Eigenschaften nicht von Wasser unterscheidet, und Hopfen und Malz gebraut wird. Diesem Szenario begegnet Horgan mit folgender Fragestellung: „(...) but why on earth should what happens in such Twin Earthly environments be considered pertinent to questions about the quausal relevance of properties instantiated by cause-events in our earthly environment?“³⁴³ Die Frage nach der quausalen Relevanz von Eigenschaften in der aktuellen Welt bleibe also von solchen möglichen Welten schlichtweg unberührt. Schließlich unterscheiden diese sich so gravierend von der aktuellen Welt, dass sie gewiss nicht als PSW gehandelt werden können. Sollte jemand Gegenteiliges behaupten, so lade er damit die Beweislast auf sich, so Horgan, weil es dann nämlich nachzuweisen gilt, warum solche möglichen Welten entgegen aller Intuition doch als PSW gelten sollten. Die zweite Strategie eine mögliche Welt zu imaginieren, in der ein Ereignis eintritt, das dem in der aktuellen Welt angemessen ähnlich und das gleichzeitig nicht die Eigenschaft ein Verlangen nach Bier zu sein instanziiert, ist *kopf-intern*. In diesen Szenarien bleibt zwar das physikalische Ereignis in Freds Gehirn, das in der aktuellen Welt ein Verlangen nach einem Bier ist, genauso bestehen wie es in der aktuellen Welt ist. Allerdings wird in solchen möglichen Welten die neuronale Verdrahtung in einem so dramatischen Ausmaß verändert, dass in diesem neuen Rahmen physikalische Ereignisse die in der aktuellen Welt noch ein Verlangen-nach-Bier-Token waren im Kausalgefüge dieser möglichen Welten nunmehr keine Instanziierung der Eigenschaft ein Verlangen nach einem Bier zu sein mehr sind. Nach Horgan können auch solche möglichen Welten nicht unter die Kategorie der PSWs subsumiert werden. Denn die neuronale Verdrahtung des Ereignisses, das in der aktuellen Welt Freds Verlangen nach einem Bier war, ist eine wichtige Hintergrundbedingung für die Verursachung

³⁴² Vgl. Hilary Putnams Gedankenexperiment mit der Zwillingerde, das später im Abschnitt über den Externalismus noch ausführlich dargestellt wird.

³⁴³ Horgan, Terence: „*Mental Quasation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 63.

von Freds Handlung und bleibt daher sicherlich über alle PSWs, die sich mit dieser Kausalabfolge beschäftigen, konstant.³⁴⁴

Da sowohl die kopf-interne als auch die kopf-externe Strategie somit laut Horgan zum Scheitern verurteilt ist, kommt er zu folgendem, abschließenden Fazit über die quausale Relevanz mentaler Eigenschaften:

It appears, therefore, that the mental properties of mental events are indeed explanatorily relevant in the causal transactions in which those events are causes. Even though we humans evidently are psycho-chemical systems all of whose behavior and inner goings-on are explainable, in principle, in terms of laws of physics as applied to our microphysical parts, we are also creatures whose mental states, qua mental, are causally efficacious. It is nice to have philosophical justification for believing in mental quausation, since most of us – myself included – could no more believe otherwise than we could fly.³⁴⁵

Schlussendlich scheinen also die (mentalen) Eigenschaften von mentalen Ereignissen tatsächlich explanatorisch relevant zu sein – und zwar in den Kausalabfolgen, in denen diese Ereignisse als Ursachen figurieren. Dies ist durchaus damit verträglich, dass Menschen psycho-chemische Systeme sind, deren gesamtes Verhalten und alle inneren Vorgänge – zumindest prinzipiell – durch die Anwendung der Naturgesetze auf mikrophysikalische Partikel erklärbar sind, so Horgan. Andererseits gilt für Menschen gleichzeitig aber auch, dass ihre mentalen Zuständen qua Mentalem, also aufgrund deren Zugehörigkeit zu einer mentalen Eigenschaft, kausal wirksam sind.

³⁴⁴ vgl. Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 62f.

³⁴⁵ Horgan, Terence: „*Mental Quausation*“ in: *Philosophical Perspectives*, Vol. 3, 1989, S. 64.

Externalismus und Anti-Individualismus als Problem für mentale Verursachung

Exkurs: Externalismus und Anti-Individualismus

Externalismus – Hilary Putnam

Als Ausgangspunkt für seine Formulierung des Externalismus dient Hilary Putnam in *The meaning of 'meaning'* der gravierende Unterschied zwischen dem zeitgenössischen wissenschaftlichen Kenntnisstand der Syntax und demjenigen der Semantik. Während auf dem Gebiet der Syntax bisweilen große Fortschritte erlangt wurden, sind die Errungenschaften im Bereich der Semantik weitaus dürftiger, so Putnam. Dies wiederum hängt Putnam zufolge insbesondere mit dem vorwissenschaftlichen Verständnis von ‚Bedeutung‘ zusammen, welches im Vergleich zu demjenigen der Syntax unterentwickelt sei. Grundsätzlich gibt es eine gewisse Ambiguität bei der Verwendung des Begriffs ‚Bedeutung‘, die für gewöhnlich in den beiden Dimensionen ‚Intension‘ und ‚Extension‘³⁴⁶ ihren Niederschlag findet. Wobei die Extension eines Begriffs schlichtweg die Menge aller Dinge ist, auf die der Begriff zutrifft bzw. von denen er wahr ist. So ist die Extension des Begriffs ‚Hase‘ in seiner Standardverwendung im Deutschen die Menge aller Hasen. Allerdings gibt es Begriffe die zwar dieselbe Extension besitzen, von denen man jedoch mitnichten behaupten würde sie hätten die gleiche Bedeutung – z.B. ‚Lebewesen mit Herz‘ und ‚Lebewesen mit Niere‘.³⁴⁷ An dieser Stelle kommt die zweite Dimension des Begriffs ‚Bedeutung‘ ins Spiel, die Intension. Darunter versteht man in etwa das Konzept, das hinter einem Begriff steht oder die Liste (oder zumindest eine Teilmenge daraus) derjenigen Eigenschaften und Merkmale, die eine Entität notwendigerweise aufweisen muss, um unter den in Frage stehenden Begriff zu fallen. Dieses Verständnis von ‚Bedeutung‘ hat zweierlei zur Folge: erstens die Auffassung Bedeutungen seien mentale Entitäten, weil der intensionale Aspekt von ‚Bedeutung‘ identisch mit dem Konzept hinter einem Begriff ist und diesem Konzept eine mentale Natur zukommt. Frege und Carnap lehnen einen

³⁴⁶ Gottlob Frege verwendet hierfür die Begriffe ‚Sinn‘ und ‚Bedeutung‘. Da jedoch in der umgangssprachlichen Verwendung der deutschen Sprache Freges Begriff ‚Sinn‘ meistens mit ‚Bedeutung‘ wiedergegeben wird, erscheint mir die Verwendung der Begriffe ‚Intension‘ und ‚Extension‘ unverfänglicher.

³⁴⁷ Streng genommen handelt es sich bei der Wendung ‚Lebewesen mit Herz‘ natürlich nicht um einen einzelnen Begriff. Allerdings lassen sich natürlich auch bei einzelnen Begriffen Beispiele finden, die zwar dieselbe Extension besitzen, aber nicht die gleiche Bedeutung, so etwa Freges klassisches Exempel ‚Morgenstern‘, ‚Abendstern‘ und ‚Venus‘.

solchen ‚Psychologismus‘ jedoch ab und fordern stattdessen, Bedeutungen als ein öffentlich zugängliches Gut zu begreifen, weil ein und dieselbe Bedeutung im gleichen Augenblick von mehr als einer Person oder von mehreren Personen zu unterschiedlichen Zeitpunkten erfasst werden kann. Diese Sichtweise impliziert den Appell, Konzepte als abstrakte Entitäten aufzufassen. Allerdings ist der einzelne Akt des Erfassens der Bedeutung selbst unter dieser Prämisse ein individueller, psychologischer Vorgang. Gemäß dieser Betrachtungsweise ist die Kenntnis der Intension eines Wortes also gleichbedeutend damit sich in einem bestimmten psychologischen Zustand zu befinden. Die zweite Konsequenz, die sich aus dem traditionellen Verständnis von ‚Bedeutung‘ ergibt ist die Annahme es sei möglich, dass zwei Begriffe mit derselben Extension und verschiedenen Intensionen existieren können, der umgekehrte Fall jedoch unmöglich sei.³⁴⁸ Insgesamt beruht die ‚klassische‘ Bedeutungstheorie nach Putnam im Wesentlichen auf zwei Thesen des logischen Positivismus, die insbesondere von Frege und Carnap verfochten wurden. Den Wortlaut dieser Thesen und seinen Standpunkt diesbezüglich gibt Putnam folgendermaßen wider:

- (I) That knowing the meaning of a term is just a matter of being in a certain psychological state (in the sense of ‚psychological state‘, in which states of memory and psychological dispositions are ‚psychological states‘; no one thought that knowing the meaning of a word was a continuous state of consciousness of course).
- (II) That the meaning of a term (in the sense of ‚intension‘) determines its extension (in the sense that the sameness of intension entails sameness of extension).

I shall argue that these two assumptions are not jointly satisfied by *any* notion, let alone any notion of meaning. The traditional concept of meaning is a concept which rests on a false theory.³⁴⁹

Die erste These besagt also, dass die Kenntnis der Bedeutung eines Begriffs gleichbedeutend damit ist sich in einem bestimmten psychologischen Zustand zu befinden. Die zweite Behauptung erklärt, dass die Bedeutung – im Sinne von Intension – eines Begriffs seine Extension festlegt. D.h., wenn zwei Begriffe dieselbe Intension haben, dann besitzen sie auch zwangsläufig die gleiche Begriffsextension. Putnam allerdings ist der Auffassung, dass diese beiden Thesen von keinem einzigen Begriff gleichzeitig erfüllt werden können. Aus der gemeinsamen Annahme von (I) und (II) folgt, dass es für zwei Individuen unmöglich ist sich in ein und demselben psychologischen Zustand zu befinden und dennoch einen Begriff A unterschiedlich zu verstehen, so Putnam. Denn der psychologische Zustand eines Sprechers determiniere die Intension von A und nach (II) somit auch dessen Extension. Genau diese Schlussfolgerung aus den beiden oben genannten Thesen hält Putnam jedoch für falsch – gemäß seiner Theorie

³⁴⁸ vgl. Putnam, Hilary: *Philosophical Papers Volume 2: Mind, Language and Reality*, Cambridge, Cambridge University Press, 1975, S. 215ff.

³⁴⁹ Putnam, Hilary: *Philosophical Papers Volume 2: Mind, Language and Reality*, Cambridge, Cambridge University Press, 1975, S. 219.

ist es nämlich durchaus möglich, dass sich zwei Sprecher in exakt dem gleichen psychologischen Zustand befinden, obwohl sich die Extension des Begriffs A in der Sprache des Einen von der Extension von A im Ideolekt des Anderen unterscheidet. In anderen Worten behauptet Putnam, dass die Extension eines Begriffs *nicht* durch den psychologischen Zustand eines Sprechers determiniert wird. Zur Untermauerung dieser These entwirft Putnam folgendes Gedankenexperiment:

(...) [S]uppose that somewhere in the galaxy there is a planet we shall call Twin Earth. Twin Earth is very much like Earth; in fact, people on Twin Earth even speak *English*. In fact, apart from the differences we shall specify in our science-fiction examples, the reader may suppose that Twin Earth is *exactly* like Earth. He may even suppose that he has a *Doppelgänger* – an identical copy – on Twin Earth, if he wishes, although my stories will not depend on this.

Although some of the people on Twin Earth (say, the ones who call themselves ‘Americans’ and the ones who call themselves ‘Canadians’ and the ones who call themselves ‘Englishmen’, etc.) speak English, there are, not surprisingly, a few tiny differences which we will now describe between the dialects of English spoken on Twin Earth and Standard English. These differences themselves depend on some of the peculiarities of Twin Earth.

One of the peculiarities of Twin Earth is that the liquid called ‘water’ is not H₂O but a different liquid whose chemical formula is very long and complicated. I shall abbreviate this chemical formula simply as XYZ. I shall suppose that XYZ is indistinguishable from water at normal temperatures and pressures. In particular, it tastes like water and it quenches thirst like water. Also, I shall suppose that the oceans and lakes and seas of Twin Earth contain XYZ and not water, that it rains XYZ on Twin Earth and not water, etc.

If a spaceship from Earth ever visits Twin Earth, then the supposition at first will be that ‘water’ has the same meaning on Earth and on Twin Earth. This supposition will be corrected when it is discovered that ‘water’ on Twin Earth is XYZ, and the Earthian spaceship will report somewhat as follows:

‘On Twin Earth the word “water” means XYZ.’

(...)

Symmetrically, if a spaceship from Twin Earth ever visits Earth, the supposition at first will be that the word ‘water’ has the same meaning on Twin Earth and on Earth. This supposition will be corrected when it is discovered that ‘water’ on Earth is H₂O, and the Twin Earthian spaceship will report:

‘On Earth the word “water” means H₂O.’

Note that there is no problem about the extension of the term ‘water’. The word simply has two different meanings (as we say): in the sense in which it is used on Twin Earth, the sense of water_{TE}, what *we* call ‘water’ simply isn’t water; while in the sense in which it is used on Earth, the sense of water_E, what the Twin Earthians call ‘water’ simply isn’t water.³⁵⁰

Das Gedankenexperiment Putnams nimmt seinen Ausgang an einem Planeten namens Zwillingserde, im Folgenden kurz ‚Zwerde‘ genannt. Die Zwerde ist ein nahezu exaktes Duplikat unserer Erde. Dementsprechend gibt es auf der Zwerde auch eine Sprachgemeinschaft, die Deutsch spricht. Da die Zwerde in ihrem strukturellen Aufbau jedoch gewisse Unterschiede zur Erde aufweist, gibt es einige kleine Unterschiede zwischen dem Deutsch, das auf der Zwerde gesprochen wird und dem irdischen Standarddeutsch. Ein solcher Unterschied bezieht sich auf die Flüssigkeit, die gemeinhin mit ‚Wasser‘ bezeichnet wird. Während das Wasser auf der Erde die molekulare Struktur H₂O besitzt, ist das Wasser der Zwerde aus anderen

³⁵⁰ Putnam, Hilary: *Philosophical Papers Volume 2: Mind, Language and Reality*, Cambridge, Cambridge University Press, 1975, S. 223f.

Elementarteilchen aufgebaut, deren Darstellung einer sehr langen und komplizierten chemischen Formel bedarf, die fortan abgekürzt als XYZ wiedergegeben wird. Angenommen, XYZ ist bei normalen Temperatur- und Druckverhältnissen ununterscheidbar von Wasser – d.h. es sieht aus wie Wasser, schmeckt wie Wasser, löscht Durst wie Wasser, etc. Zudem seien die Flüsse der Zwerde nicht mit Wasser, sondern mit XYZ gefüllt, genau wie die Seen und Meere. Bei Regen fällt auf der Zwerde XYZ vom Himmel, usw. Sollte nun eines Tages ein Raumschiff von der Erde auf der Zwerde landen, so ist anzunehmen, dass dessen Insassen zunächst glauben würden, dass das Wort ‚Wasser‘ auf der Zwerde dieselbe Bedeutung hat wie auf der Erde. Diese Vermutung würde allerdings korrigiert werden müssen, sobald man entdeckt, dass das, was auf der Zwerde ‚Wasser‘ genannt wird, die Mikrostruktur XYZ aufweist. Vermutlich würden die Wissenschaftler an Bord des Raumschiffs so etwas wie ‚Auf der Zwerde bedeutet das Wort „Wasser“ XYZ.‘ berichten. Ganz analog würden auch zirdische Kosmonauten melden, dass der Begriff „Wasser“ auf der Erde H_2O bedeutet. Laut Putnam ist zu beachten, dass es hier keine Schwierigkeiten hinsichtlich der Extension des Begriffs ‚Wasser‘ gibt, sondern, dass das Wort vielmehr zwei unterschiedliche Bedeutungen bzw. Intensionen besitzt. In seiner zirdischen Gebrauchweise ($Wasser_Z$) ist das, was wir ‚Wasser‘ nennen ($Wasser_E$) kein Wasser und umgekehrt. Die Extension von $Wasser_Z$ ist in etwa die Menge aller Entitäten, die aus XYZ-Molekülen bestehen, wohingegen sich $Wasser_E$ grob gesagt auf alle Objekte bezieht, die aus H_2O -Molekülen zusammengesetzt sind. Bislang scheint Putnams Gedankenexperiment noch unverfänglich zu sein; allerdings lädt er den Leser nun zu einer Zeitreise ein.

Now let us roll back time to about 1750. At that time chemistry was not developed on either Earth or Twin Earth. The typical Earthian speaker of English did not know that water consisted of hydrogen and oxygen, and the typical Twin Earthian speaker of English did not know ‘water’ consisted of XYZ. Let Oscar₁ be such a typical Earthian English speaker, and let Oscar₂ be his counterpart on Twin Earth. You may suppose that there is no belief that Oscar₁ had about water that Oscar₂ did not have about ‘water’. If you like, you may even suppose that Oscar₁ and Oscar₂ were exact duplicates in appearance, feelings, thoughts, interior monologue, etc. Yet the extension of the term ‘water’ was just as much H_2O on Earth in 1750 as it is in 1950; and the extension of the term ‘water’ was just as much XYZ on Twin Earth in 1750 as in 1950. Oscar₁ and Oscar₂ understood the term ‘water’ differently in 1750 *although they were in the same psychological state*, and although, given the state of science at the time, it would have taken their scientific communities about fifty years to discover that they understood the term ‘water’ differently. Thus the extension of the term ‘water’ (and, in fact, its ‘meaning’ in the intuitive preanalytical usage of the term) is *not* a function of the psychological state of the speaker by itself.³⁵¹

³⁵¹ Putnam, Hilary: *Philosophical Papers Volume 2: Mind, Language and Reality*, Cambridge, Cambridge University Press, 1975, S. 224.

Im Jahr 1750 wussten weder die Zwerden- noch die Erdenbewohner über die molekulare Struktur von Wasser und Zwasser³⁵² Bescheid. Angenommen, Oskar₁ ist ein typischer irdischer Sprecher des Deutschen zu jener Zeit und Oskar₂ sein zirdisches Pendant. Zudem habe Oskar₁ keine Annahme über Wasser, die Oskar₂ nicht auch im Bezug auf Zwasser hat. Schließlich verhalten sich Oskar₁ und Oskar₂ hinsichtlich ihrer gesamten Erscheinung, all ihren Gefühlen, Gedanken, inneren Monologe, usw. wie zwei identische Kopien. Allerdings war die Extension des Begriffs ‚Wasser‘ 1750 genau die gleiche wie 1950 – d.h. auf der Erde war die Extension von ‚Wasser‘ 1750 genau wie 1950 H₂O, während sie auf der Zwerde zu beiden Zeitpunkten XYZ war. Daher haben Oskar₁ und Oskar₂ 1750 den Terminus ‚Wasser‘ auf *unterschiedliche* Art und Weise verstanden, *obwohl* sich beide in exakt dem *gleichen* psychologischen Zustand befunden haben. Folglich kann die Extension des Wortes ‚Wasser‘ nicht allein eine Funktion des psychologischen Zustands des Sprechers sein. Da jedoch aus der gemeinsamen Annahme von (I) und (II) wie oben beschrieben folgt, dass der psychologische Zustand eines Sprechers die Extension eines von ihm verwendeten Begriffes festlegt, beweist das Gedankenexperiment laut Putnam, dass (I) und (II) nicht beide zugleich wahr sein können. Genau dies wiederum wird ihm zufolge allerdings vom traditionellen Verständnis von ‚Bedeutung‘ behauptet. Folgt man Putnams Gedankenexperiment, so hängt die Bedeutung – im Sinne von Extension – eines Begriffs nicht (ausschließlich) von den internen Zuständen des Sprechers ab, der das Wort benutzt, sondern auch von der Beschaffenheit der Welt.

Zur weiteren Illustration seines Kerngedanken nimmt Putnam eine kleine Modifikation seines Gedankenexperiments vor:

Let us now modify our science-fiction story. I do not know whether one can make pots and pans out of molybdenum; and if one can make them out of molybdenum, I don't know whether they could be distinguished easily from aluminium pots and pans. (I don't know any of this even though I have acquired the word 'molybdenum'.) So I shall suppose that molybdenum pots and pans *can't* be distinguished from aluminium pots and pans save by an expert. (To emphasize the point, I repeat that this could be true for all I know, and *a fortiori* it could be true for all I know by virtue of 'knowing the meaning' of the words *aluminium* and *molybdenum*.) We will now suppose that molybdenum is as common on Twin Earth as aluminium is on Earth, and that aluminium is as rare on Twin Earth as molybdenum is on Earth. In particular, we shall assume that 'aluminium' pots and pans are made of molybdenum on Twin Earth. Finally, we shall assume that the words 'aluminium' and 'molybdenum' are *switched* on Twin Earth: 'aluminium' is the name of *molybdenum* and 'molybdenum' is the name of *aluminium*. This example shares some features with the previous one. (...) But there is one important difference between the two cases. An Earthian metallurgist could tell very easily that 'aluminium' was molybdenum, and a Twin Earthian metallurgist could tell equally easily that aluminium was 'molybdenum'. (...) Whereas in 1750 no one on either Earth or Twin Earth could have distinguished water from 'water', the confusion of aluminium and 'aluminium' involves only a part of the linguistic communities involved. The example makes the same point as the preceding one. If Oscar₁ and Oscar₂ are standard speakers of Earthian English and Twin Earthian English respectively, and neither is chemically or metallurgically sophisticated, then

³⁵² Zur einfacheren Unterscheidung verwende ich im Folgenden den Begriff ‚Zwasser‘ für das, was auf der Zwerde mit ‚Wasser‘ bezeichnet wird.

there may be no difference at all in their psychological state when they use the word ‘aluminium’; nevertheless we have to say that ‘aluminium’ has the extension of *aluminium* in the idiolect of Oscar₁ and the extension of *molybdenum* in the idiolect of Oscar₂. (Also we have to say that Oscar₁ and Oscar₂ mean different things by ‘aluminium’, that ‘aluminium’ has a different meaning on Earth than it does on Twin Earth, etc.) Again we see that the psychological state of the speaker does *not* determine the extension (*or* the ‘meaning’, speaking preanalytically) of the word.³⁵³

Angenommen, man könnte Pfannen und Töpfe aus Molybdän herstellen und nur Fachmänner könnten diese von herkömmlichem Aluminiumkochgeschirr unterscheiden. Zudem seien die Aluminiumvorkommnisse auf der Zwerde so selten wie die Molybdänressourcen auf der Erde und Molybdän sei auf der Zwerde im selben Ausmaß vorhanden wie Aluminium auf der Erde. Schließlich gehe man noch davon aus, dass die Begriffe ‚Aluminium‘ und ‚Molybdän‘ auf der Zwerde vertauscht sind: ‚Aluminium‘ (im Folgenden Zwaluminium) bedeutet auf der Zwerde also Molybdän und mit ‚Molybdän‘ (fortan Zwmolybdän) nehmen Zwerdlinge auf Aluminium Bezug. Trotz der vielen Parallelen zum vorhergehenden Gedankenexperiment gibt es einen wesentlichen Unterschied: ein Experte im Bereich der Metallurgie könnte schnell erkennen, dass es sich bei Zwaluminium um Molybdän und bei Zwmolybdän um Aluminium handelt. Anders als beim Wasser-Beispiel im Jahr 1750, betrifft die Verwechslung von Aluminium und Zwaluminium also nur einen bestimmten Personenkreis aus der gesamten Sprachgemeinschaft. Allerdings demonstriert auch diese Spielart des Gedankenexperiments, dass weder die Extension eines Begriffs, noch dessen Bedeutung im vorwissenschaftlichen Sinn ausschließlich von dem psychologischen Zustand des Sprechers festgelegt werden.

Abschließend präsentiert Putnam noch folgende Variante seines Gedankenexperiments:

Suppose you are like me and cannot tell an elm from a beech tree. We will still say that the extension of ‘elm’ in my idiolect is the same as the extension of ‘elm’ in anyone else’s, viz., the set of all elm trees, and that the set of all beech trees is the extension of ‘beech’ in *both* of our idiolects (as it should). Is it really credible that this difference in extension is brought about by some difference in our *concepts*? My *concept* of an elm tree is exactly the same as my concept of a beech tree (I blush to confess). (This shows that the identification of meaning ‘in the sense of intension’ with *concept* cannot be correct, by the way.) If someone heroically attempts to maintain that the difference between the extension of ‘elm’ and the extension of ‘beech’ in *my* idiolect is explained by a difference in my psychological state, then we can always refute him by constructing a ‘Twin Earth’ example – just let the words ‘elm’ and ‘beech’ be switched on Twin Earth (the way ‘aluminium’ and ‘molybdenum’ were in the previous example). Moreover, I suppose I have a *Doppelgänger* on Twin Earth who is molecule for molecule ‘identical’ with me (in the sense in which two neckties can be ‘identical’). If you are a dualist, then also suppose that my *Doppelgänger* thinks the same verbalized thoughts I do, has the same sense data, the same dispositions, etc. It is absurd to think *his* psychological state is one bit different from mine: yet he ‘means’ *beech* when he says ‘elm’ and I ‘mean’ *elm* when I say elm. Cut the pie any way you like, ‘meanings’ just ain’t in the *head*!³⁵⁴

³⁵³ Putnam, Hilary: *Philosophical Papers Volume 2: Mind, Language and Reality*, Cambridge, Cambridge University Press, 1975, S. 225f.

³⁵⁴ Putnam, Hilary: *Philosophical Papers Volume 2: Mind, Language and Reality*, Cambridge, Cambridge University Press, 1975, S. 226f.

Sicherlich gibt es in der deutschen Sprachgemeinschaft einige Menschen, die Ulmen nicht von Buchen zu unterscheiden wissen. Dennoch würde man natürlich sagen, dass auch in deren Verwendung des Deutschen das Wort ‚Ulme‘ dieselbe Extension hat wie im Vokabular derjenigen, die in der Lage sind Ulmen und Buchen auseinanderzuhalten – nämlich die Menge aller Ulmen (ganz analog ist auch die Extension des Wortes ‚Buche‘ die Menge aller Buchen im Sprachgebrauch beider Gruppen). Aus diesem Grund erscheint es unglaublich, dass der Unterschied in Bezug auf die Extension der beiden Begriffe durch die zugrundeliegenden Konzepte zu Stande kommen könnte. Schließlich sind die besagten Konzepte bei einer bestimmten Gruppe von Deutschsprachlern deckungsgleich. Sollte man nun argumentieren, dass bei einer Person, die Buchen und Ulmen nicht voneinander unterscheiden kann, die Unterschiede hinsichtlich der Extension dieser beiden Begriffe im Sprachgebrauch dieses Menschen, sich durch Unterschiede in seinem psychologischen Zustand erklären lassen, so ließe sich sehr einfach ein Gegenbeispiel entwickeln. Angenommen, auf der Zwerde sind die Begriffe ‚Ulme‘ und ‚Buche‘ vertauscht und gleichzeitig besitzt die oben genannte Person dort einen Doppelgänger, dem sie bis aufs letzte Atom gleicht, so wäre es absurd anzunehmen, dass sich die psychologischen Zustände der beiden auch nur im geringsten voneinander unterscheiden. Nichtsdestotrotz bezieht sich der Erdling auf Buchen, wenn er von ‚Buchen‘ spricht, während der Zwerdling hingegen auf Ulmen referiert, wenn er von ‚Buchen‘ spricht. Darin liegt nach Putnam der endgültige Beweis, dass Bedeutungen sozusagen nicht im Kopf sein können, sondern immer ein kontextabhängiges Element besitzen.

Anti-Individualismus – Tyler Burge

Im Diskurs hinsichtlich der Philosophie des Geistes macht Tyler Burge seit Hegels *Phänomenologie des Geistes* eine gewisse Janusköpfigkeit aus: einerseits sei man am Individuum als Träger von mentalen Zuständen und Ereignissen interessiert, andererseits wurden aber auch soziale Institutionen in den Fokus gerückt, weil jene den Einzelnen und den Inhalt seiner Gedanken maßgeblich prägen. Allerdings wurden den sozialen Einflüssen im Vergleich zu den individuellen Aspekten im Gedankenaustausch der Philosophen weitaus weniger nüchterne Aufmerksamkeit eingeräumt, so Burge – jedoch völlig zu Unrecht und deshalb proklamiert er: „I shall offer some considerations that stress social factors in descriptions of an individual’s mental phenomena. These considerations call into question individualistic presuppositions of

several traditional and modern treatments of mind.³⁵⁵ Laut Burge spielen soziale Faktoren also selbst bei der Beschreibung mentaler Phänomene auf der individuellen Ebene eine entscheidende Rolle. Daher seien rein individualistisch orientierte Theorien des Geistes stets mit gewissen Mängeln behaftet.

An den Beginn seiner Untersuchung stellt Burge zunächst eine Unterscheidung zweier Arten über Mentales zu sprechen. Während die eine Variante mittels satzartiger Ausdrücke auf mentale Zustände oder Ereignisse referiert, verzichtet die andere auf solche Konstrukte. Ein Beispiel für die erste Art der Bezugnahme ist ‚Alfred denkt, dass das Sofa seiner Freunde hässlich ist‘. Im Gegensatz dazu ist ‚Alfred verspürt Schmerzen‘ ein archetypisches Exempel für die zweite Art des Rekurses auf Mentales. Typischerweise werden Gedanken, Glaubenssätze und Intentionen mithilfe von ‚dass-Sätzen‘ ausgedrückt, während diese Ausdrucksform bei Gefühlen und Schmerzen hingegen in der Regel nicht gewählt wird. Genauer gesagt wird mit den ‚dass-Sätzen‘ der *Inhalt* eines mentalen Zustands oder Ereignisses spezifiziert. So drückt z.B. die Phrase ‚dass Sofas bequemer sind als Kirchenbänke‘ den Inhalt von Alfreds Glauben aus, dass Sofas bequemer als Kirchenbänke sind. Die Frage wann die Inhalte von Gedanken, Glaubenssätzen, etc. unterschiedlich bzw. gleich sind, ist maßgeblich abhängig von der zugrundeliegenden Theorie mentalistischer Sprache. Allerdings gebe es aber auch völlig unkontroverse Fälle, so Burge: „On any systematic theory, differences in the *extension* – the actual denotation, referent or application – of counterpart expressions in that-clauses will be semantically represented, and will, in our terms, make for differences in content.“³⁵⁶ Sollten also zwei Wendungen in ‚dass-Sätzen‘, die Gedanken ausdrücken, verschiedene Extensionen besitzen, so haben diese Gedanken Burge zufolge notwendigerweise einen unterschiedlichen Inhalt. Allerdings müsse man sich hüten daraus den Umkehrschluss zu ziehen - nämlich, dass extensionsgleiche Termini in solchen ‚dass-Sätzen‘ *salva veritate* miteinander austauschbar wären. Dazu Burge:

Thus from the facts that water is H₂O and that Bertrand thought that water is not fit to drink, it does not follow that Bertrand thought that H₂O is not fit to drink. When an expression like ‘water’ functions in a content clause so that it is not freely exchangeable with all extensionally equivalent expressions, we shall say that it has *oblique occurrence*. Roughly speaking, the reason why ‘water’ and ‘H₂O’ are not interchangeable in our report of Bertrand’s thoughts is that ‘water’ plays a role in characterizing a different mental act or state from that which ‘H₂O’ would play in characterizing. In this context at least, thinking that water is not fit to drink is different from thinking that H₂O is not fit to drink.³⁵⁷

³⁵⁵ Burge, Tyler: „*Individualism and the Mental*“ in: *Midwest Studies in Philosophy*, Vol. 4, 1979, S. 74.

³⁵⁶ Burge, Tyler: „*Individualism and the Mental*“ in: *Midwest Studies in Philosophy*, Vol. 4, 1979, S. 75.

³⁵⁷ Burge, Tyler: „*Individualism and the Mental*“ in: *Midwest Studies in Philosophy*, Vol. 4, 1979, S. 76.

Aus den Tatsachen, dass Wasser H₂O ist und dass Bertrand denkt, dass Wasser nicht trinkbar ist, folgt gemäß Burge nicht (zwangsläufig), dass Bertrand der Auffassung ist, dass H₂O nicht trinkbar ist. Der Grund weshalb ‚Wasser‘ in der Rede über Bertrands Gedanken nicht durch ‚H₂O‘ ersetzt werden darf, besteht darin, dass mit ‚Wasser‘ auf einen anderen mentalen Akt bzw. Zustand referiert wird als mit ‚H₂O‘. Der Gedanke, dass Wasser nicht trinkbar ist unterscheidet sich ergo von dem Gedanken, dass H₂O nicht trinkbar ist. Wenn ein Wort in ‚dass-Sätzen‘, die Gedanken zum Ausdruck bringen, nicht durch einen beliebigen extensionsgleichen Ausdruck ersetzbar ist, so bezeichnet Burge dies als ein obliques bzw. schiefes Vorkommen des entsprechenden Begriffs. Diese Art der Verwendung von Begriffen kennzeichnet nach Burge eine bestimmte Art mentaler Rede: „Mentalistic discourse containing obliquely occurring expressions has traditionally been *intentional discourse*.“³⁵⁸ Wenn Ausdrücke im Diskurs über Mentales also schief verwendet werden, so handelt es sich um intentionale Redeweise.

Vor diesem gedanklichen Hintergrund stellt Burge ein dreistufiges Gedankenexperiment vor:

We now turn to a three-step thought experiment. Suppose first that:

A given person has a large number of attitudes commonly attributed with content clauses containing ‘arthritis’ in oblique occurrence. For example, he thinks (correctly) that he has had arthritis for years, that his arthritis in his wrists and fingers is more painful than his arthritis in his ankles, that it is better to have arthritis than cancer of the liver, that stiffening joints is a symptom of arthritis, that certain sorts of aches are characteristic of arthritis, that there are various kinds of arthritis, and so forth. In short, he has a wide range of such attitudes. In addition to these unsurprising attitudes, he thinks falsely that he has developed arthritis in the thigh.

Generally competent in English, rational and intelligent, the patient reports to his doctor his fear that his arthritis has now lodged in his thigh. The doctor replies by telling him that his cannot be so, since arthritis is specifically an inflammation of joints. Any dictionary could have told him the same. The patient is surprised, but relinquishes his view and goes on to ask what might be wrong with his thigh.³⁵⁹

Der erste Schritt des Gedankenexperiments besteht in der Annahme, eine Person (im Folgenden Hans) habe eine große Anzahl von Überzeugungen hinsichtlich Arthritis, die ihr anhand von Inhalts- oder dass-Sätzen zugeschrieben werden, in denen ‚Arthritis‘ schief verwendet wird. So nimmt Hans z.B. richtigerweise an, dass er schon seit Jahren unter Arthritis leidet, dass die Arthritis in seinen Fingern und Handgelenken schmerzhafter ist als die in seinen Fußgelenken, dass es besser ist an Arthritis erkrankt zu sein als an Leberkrebs, dass die schleichende Versteifung von Gelenken ein Symptom von Arthritis ist, dass bestimmte Arten von Schmerzen typisch für Arthritis sind, dass es unterschiedliche Formen der Arthritis gibt, usw. Kurzum, Hans besitzt also eine beachtliche Menge solcher Überzeugungen, allerdings vertritt er zudem die (falsche) Ansicht, dass neuerdings auch sein Oberschenkel von Arthritis

³⁵⁸ Burge, Tyler: „*Individualism and the Mental*“ in: *Midwest Studies in Philosophy*, Vol. 4, 1979, S. 76.

³⁵⁹ Burge, Tyler: „*Individualism and the Mental*“ in: *Midwest Studies in Philosophy*, Vol. 4, 1979, S. 77.

befallen wurde. Grundsätzlich nehmen wir an, Hans ist ein kompetenter Sprecher des Deutschen, handelt in der Regel rational und verfügt über normale Intelligenz. Als er nun dem behandelnden Arzt von seiner Sorge berichtet, dass sich die Arthritis bei ihm zuletzt im Oberschenkel festgesetzt habe, antwortet ihm dieser, dass dies unmöglich ist, weil Arthritis eine entzündliche Erkrankung ist, von der ausschließlich Gelenke betroffen sind. Auch ein Blick ins Lexikon hätte Hans belehrt, dass es sich bei Arthritis um eine entzündliche Gelenkerkrankung handelt. Angesichts dieser neuen Information wird Hans seine Auffassung, dass er Arthritis im Oberschenkel hat, fallenlassen und sich stattdessen beim Arzt erkundigen, woher die Schmerzen in seinem Oberschenkel ansonsten rühren könnten.

Nach dieser unstrittigen Darstellung stellt Burge den zweiten Schritt seines Gedankenexperiments vor:

The second step of the thought experiment consists of a counterfactual supposition. (...) The *counterfactuality* in the supposition touches only the patient's social environment. In actual fact, 'arthritis', as used in his community, does not apply to ailments outside joints. (...) But in our imagined case, physicians, lexicographers, and informed laymen apply 'arthritis' not only to arthritis but to various other rheumatoid ailments. (...) To summarize the second step:

The person might have had the same physical history and non-intentional mental phenomena while the word 'arthritis' was conventionally applied, and defined to apply, to various rheumatoid ailments, including the one in the person's thigh, as well as to arthritis.³⁶⁰

Die zweite Stufe von Burges Gedankenexperiment besteht also hauptsächlich in einer kontrafaktischen Annahme. Wobei sich die Kontrafaktizität ausschließlich auf Hans' soziales Umfeld bezieht. In der aktuellen Welt wird der Begriff ‚Arthritis‘ von der deutschen Sprachgemeinschaft so verwendet, dass er nicht auf ein Krankheitsbild angewendet werden kann, dessen wesentliche Symptomatik Schmerzen außerhalb von Gelenken umfasst. Dies wiederum ist schlichtweg darauf zurückzuführen, dass die Standarddefinition von ‚Arthritis‘ einen solchen Gebrauch ausschließt. In Burges kontrafaktischem Szenario hingegen hat das Wort ‚Arthritis‘ einen größeren Anwendungsbereich: So benutzen dort Ärzte, Verfasser von Lexika und auch sachkundige Laien die Bezeichnung ‚Arthritis‘ sowohl in unserem Sinn (also als entzündliche Gelenkerkrankung), als auch zusätzlich noch in einem weiteren Sinn für andere rheumatische Erkrankungen. Dabei ist durchaus vorstellbar, dass das, was in der aktuellen Welt als ‚Arthritis‘ bezeichnet wird, in der kontrafaktischen Welt nicht als eine eigenständige Gruppe von Krankheiten klassifiziert wird, oder, dass es einen anderen, allerdings nicht allgemein verwendeten, Terminus gibt, um auf arthritische Erkrankungen (in unserem Sinne) zu referieren. Zusammenfassend konstatiert Burge, die von ihm entworfene kontrafaktische Si-

³⁶⁰ Burge, Tyler: „*Individualism and the Mental*“ in: *Midwest Studies in Philosophy*, Vol. 4, 1979, S. 77f.

tuation offenbare folgende Möglichkeit: so könnte Hans' gesamte physikalische Geschichte, sowie seine mentale (allerdings nicht-intentional beschriebene) Vergangenheit in der aktuellen Welt und im kontrafaktischen Szenario identisch sein. Trotz dieser Ununterscheidbarkeit von Hans' Zuständen, würde er die Bezeichnung ‚Arthritis‘ in der aktuellen Welt falsch anwenden, wenn er sich damit auf die Schmerzen in seinem Oberschenkel bezieht, wohingegen dies in der kontrafaktischen Welt eine legitime Verwendung des Begriffs darstellen würde.

Anschließend stellt Burge die dritte Stufe seines Gedankenexperiments vor:

The final step is an interpretation of the counterfactual case, or an addition to it as so far described. It is reasonable to suppose that:

In the counterfactual situation, the patient lacks some – probably *all* – of the attitudes commonly attributed with content clauses containing ‘arthritis’ in oblique occurrence. He lacks the occurrent thoughts or beliefs that he has arthritis in the thigh, that he has had arthritis for years, that stiffening joints and various sorts of aches are symptoms of arthritis, that his father had arthritis, and so on.

We suppose that in the counterfactual case we cannot correctly ascribe any content clause containing an oblique occurrence of the term ‘arthritis’. It is hard to see how the patient could have picked up the notion of arthritis. The word ‘arthritis’ in the counterfactual community does not mean *arthritis*. It does not apply only to inflammations of joints. We suppose that no other word in the patient’s repertoire means *arthritis*. ‘Arthritis’ in the counterfactual situation, differs both in dictionary definition and in extension from ‘arthritis’ as we use it. Our ascriptions of content clauses to the patient (and ascriptions within his community) would not constitute attributions of the same contents we actually attribute. For counterpart expressions in the content clauses that are actually and counterfactually ascribable are not even extensionally equivalent. However we describe the patient’s attitudes in the counterfactual situation, it will not be with a term or phrase extensionally equivalent with ‘arthritis’. So the patient’s counterfactual attitude contents differ from his actual ones.³⁶¹

Die letzte Phase von Burges dreistufigem Gedankenexperiment ist eine vertiefende Analyse und Erläuterung des bisher beschriebenen kontrafaktischen Szenarios. So ist es durchaus vernünftig anzunehmen, dass man Hans kaum eine (vermutlich sogar überhaupt keine) Überzeugung zuschreiben kann, in der ‚Arthritis‘ in schiefer Verwendung vorkommt. Daher glaubt Hans in der kontrafaktischen Situation nicht, dass er Arthritis im Oberschenkel hat, dass er seit Jahren unter Arthritis leidet, dass die Versteifung der Gelenke und bestimmte Schmerzzustände symptomatisch für Arthritis sind, dass sein Vater an Arthritis erkrankt war, usw. Da der Begriff ‚Arthritis‘ in der kontrafaktischen Sprachgemeinschaft eine andere Bedeutung hat wie in der aktuellen Welt, ist nicht ersichtlich auf welche Weise es gelingen könnte, Hans eine Überzeugung zuzuschreiben, die mittels eines Inhaltssatzes, in dem ‚Arthritis‘ schief vorkommt, beschrieben werden kann. Denn die Vokabel ‚Arthritis‘ unterscheidet sich in der kontrafaktischen Situation sowohl in Anbetracht ihrer lexikalischen Definition, als auch hinsichtlich der Extension vom Wort ‚Arthritis‘ in der aktuellen Welt. Da die beiden Begriffe nicht dieselbe Extension besitzen, können auch die Inhalte von Hans' Überzeugungen in beiden

³⁶¹ Burge, Tyler: „*Individualism and the Mental*“ in: *Midwest Studies in Philosophy*, Vol. 4, 1979, S. 78f.

Welten nicht identisch sein. Insofern muss Hans in beiden Szenarien unterschiedliche Überzeugungen besitzen.

Abschließend zieht Burge folgendes Fazit aus seinem Gedankenexperiment:

The upshot of these reflections is that the patient's mental contents differ while his entire physical and non-intentional mental histories, considered in isolation from their social context, remain the same. (We could have supposed that he dropped dead at the time he first expressed his fear to the doctor.) The differences seem to stem from differences "outside" the patient considered as an isolated physical organism, causal mechanism, or seat of consciousness. The difference in his mental contents is attributable to differences in his social environment. In sum, the patient's internal qualitative experiences, his physiological states and events, his behaviorally described stimuli and responses, his dispositions to behave, and whatever sequences of states (non-intentionally described) mediated his input and output – all these remain constant, while his attitude contents differ, even in the extensions of counterpart notions. As we observed at the outset, such differences are ordinarily taken to spell differences in mental states and events.³⁶²

Der zentrale Aspekt, den ein Vergleich von aktueller und kontrafaktischer Welt nach Burge zu Tage fördert, besteht darin, dass die Inhalte von Hans' Überzeugungen in den beiden Szenarien *unterschiedlich* sind, *obwohl* seine gesamte physikalische und nicht-intentional beschriebene mentale Vergangenheit unter Vernachlässigung der sozialen Rahmenbedingungen *identisch* ist. Daher müssen diese Unterschiede von Begebenheiten *außerhalb* von Hans herrühren, wenn man Hans als ein in sich geschlossenes, physisches System versteht. Genauer gesagt kann man die Verschiedenartigkeit von Hans' mentalen Inhalten auf Abweichungen in seinem *sozialen* Umfeld zurückführen. Schließlich sind Hans' phänomenales Erleben, seine physiologischen Zustände, die ihm präsentierten Reize und die darauffolgenden Reaktionen in seinem Verhalten, seine Neigungen, sowie alle weiteren, nicht-intentional beschriebenen Parameter in beiden Szenarien deckungsgleich. Wohingegen die Inhalte seiner Überzeugungen und die Extensionen der entsprechenden Begriffe in beiden Umfeldern divergieren. Da mentale Zustände im Allgemeinen über ihre Inhalte individuiert werden, ist es unumgänglich Hans in den beiden Szenarien unterschiedliche mentale Zustände zuzuschreiben. Insofern weisen Hans' mentale Zustände aber eine nicht eliminierbare kontextuelle Komponente auf. D.h. Hans und eine exakte physikalische Kopie von Hans könnten sich in *unterschiedlichen mentalen* Zuständen befinden, *obwohl* sie *physikalisch* betrachtet *ununterscheidbar* sind. Folglich können Hans' mentale Zustände auch nicht auf seinen physikalischen Zuständen supervenieren.

³⁶² Burge, Tyler: „Individualism and the Mental“ in: *Midwest Studies in Philosophy*, Vol. 4, 1979, S. 79.

Externalismus und Anti-Individualismus als Problem für mentale Verursachung

Die Akzeptanz von Externalismus und Anti-Individualismus hat unmittelbare Auswirkungen auf die Intelligibilität mentaler Verursachung. Wenn nämlich der Inhalt von intentionalen Zuständen extern individuiert wird, ist dieser abhängig von Faktoren *außerhalb* des Trägers der entsprechenden Intention (z.B. von kausalen, historischen oder sozialen Beziehungen zu dessen Umwelt). Somit ist der Inhalt intentionaler Zustände keine *intrinsische*, sondern eine *extrinsische* Eigenschaft. Allgemein geht man jedoch davon aus, dass bei Verursachungsrelationen ausschließlich intrinsische Merkmale eine Rolle spielen. Daher wären intentionale Zustände für kausale Beziehungen irrelevant, wenn der intentionale Gehalt eine extrinsische Eigenschaft mentaler Zustände ist.

Fred Dretske veranschaulicht das Problem vor dem Hintergrund des Externalismus und Anti-Individualismus zu einem kohärenten Verständnis mentaler Verursachung zu gelangen anhand folgender Analogie: „(...) the mind is to intentional action what money is to the behavior of a vending machine.“³⁶³ Gemäß dieses Bildes steht der menschliche Geist in dem selben Verhältnis zu intentionalen Handlungen wie Geld zum ‚Verhalten‘ eines Verkaufsautomaten. Diese Analogie ist insbesondere deshalb angebracht, weil die Individuierungsbedingungen von Geld und z.B. von Überzeugungen nahezu identisch sind, so Dretske: „(...) just what makes money money is not its intrinsic character – shape, size, and density of the coins, for example – but certain extrinsic or relational facts about these coins (the fact that they possess monetary value), so too what makes a belief a belief is not its intrinsic neurobiological character, but rather, certain extrinsic facts about it – the fact that it has a certain meaning or content, the fact that it has certain *intentional* properties.“³⁶⁴ So sind es nicht die intrinsischen Eigenschaften einer Münze, wie beispielsweise ihre Größe, ihre Dichte, ihre Form oder ihr Gewicht, die sie zu Geld bzw. zu einem legitimen Zahlungsmittel machen, sondern diese Funktion übernehmen vielmehr extrinsische oder relationale Eigenschaften der Münze, wie etwa die gesellschaftliche Konvention, dass sie einen bestimmten Geldwert besitzt. Genauso wird eine Überzeugung nicht aufgrund ihrer intrinsischen, neurobiologischen Eigenschaften zu der Überzeugung, die sie *de facto* ist, sondern vielmehr wegen gewisser extrinsischer Eigenschaften wie

³⁶³ Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 259.

³⁶⁴ Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 259.

Inhalt oder Bedeutung. Die Auswirkungen der Akzeptanz dieser Analogie auf mentale Kausalität beschreibt Dretske folgendermaßen:

If we take this analogy seriously, it suggests that beliefs, qua beliefs, are as irrelevant to animal behavior as is money, qua money, to the behavior of vending machines. Since it is facts about the shape and size of coins, not facts about their monetary value, that explain why coins cause a machine to yield its contents, the analogy, if we take it seriously – and a good many philosophers do – compels us to conclude that it is the intrinsic features of belief, their neurobiological properties, not their extrinsic properties, their meaning or content, that explains why we do what we do. We thus seem to be driven to the conclusion that what we believe is causally irrelevant to what we do.³⁶⁵

Wenn man ein Geldstück in einen Automaten wirft und dieser anschließend z.B. einen Schokoladenriegel ausgibt, so kann man davon ausgehen, dass der Automat nur auf *bestimmte* Eigenschaften der Münze reagiert hat. Genau genommen ist davon auszugehen, dass der Automat z.B. Durchmesser, Dicke, Gewicht und magnetische Permeabilität des Geldstücks prüft und abhängig von diesen Attributen die Süßware ausgibt oder eben nicht. D.h. eine gelungene Fälschung der entsprechenden Münze(n), die mit echten Münzen in den oben genannten Merkmalen übereinstimmt, würde auch zum Auswurf der Süßigkeit führen, obwohl die gefälschte Münze keinen Geldwert besitzt und somit im eigentlichen Sinne kein Geldstück ist. Die Eigenschaft der Münze Geld zu sein, wäre also für ihre kausale Kraft hinsichtlich der Ausgabe von Schokoladenriegeln aus dem Automaten vollkommen bedeutungslos. Schließlich würde auch der Einwurf einer hinlänglich gut gefälschten Münze die Ausgabe einer Süßigkeit aus dem Verkaufsautomaten verursachen.

Falls die besagte Analogie das Wesen intentionaler Eigenschaften korrekt erfasst, steht man laut Dretske vor dem Problem, dass z.B. Überzeugungen nicht qua Überzeugung, d.h. in ihrer Eigenschaft eine Überzeugung zu sein, intentionale Handlungen verursachen. Infolgedessen spielen intentionale Eigenschaften bei der Verursachung von Handlungen keine Rolle, sie sind in anderen Worten kausal irrelevant für die Produktion von Verhalten. Genauso wie bei Münzen etwa Größe und Gewicht erklären warum der Einwurf einer Münze die Ausgabe eines Schokoriegels bewirkt und nicht deren Geldwert, so sind es bei Überzeugungen auch deren neurobiologische Eigenschaften, die das resultierende Verhalten erklären, und nicht deren Inhalt. Alles im allem liegt also die Schlussfolgerung nahe, dass die Inhalte von Überzeugungen kausal irrelevant für Verhalten sind.

In seinem Aufsatz *Content, Causation, and Psychophysical Supervenience* befasst sich Joseph Owens zunächst mit der Alltagspsychologie (englisch: folk psychology), die das Verhalten

³⁶⁵ Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 259f.

eines Individuums im Allgemeinen mit Gründen erklärt. So unproblematisch dieses Verfahren auf den ersten Blick auch erscheint, lässt es bei genauerer Betrachtung jedoch zahlreiche Fragen offen. Dementsprechend skizziert Owens das Hauptanliegen seines Aufsatzes folgendermaßen:

In this paper, my concern is with the nature of the explanatory states, and the nature of the explanatory linkage. Though the debate is far from closed, there is something of a received view on both issues. The received view of the explanatory states is that these states are propositional attitudes, representational or content-individuated states. Call this the “representational thesis”. On the nature of the explanatory linkage, the popular view is that psychological explanation is causal in character. Call this the “causal thesis”. (...) I argue that this widely shared conception of folk psychological explanation – the “C/R Model” (causal/representational model) – is mistaken. I will not argue that the causal thesis is incorrect or that the representational thesis is incorrect, but rather that a serious tension lies between them, and that one cannot have both.³⁶⁶

Owens sieht seine zentrale Aufgabe also darin, das Wesen von Zuständen, die Verhalten erklären sollen, sowie die Beschaffenheit der erklärungsstiftenden Verbindung zwischen psychologischen Zuständen und Verhalten genauer zu ergründen. Zu beiden Fragestellungen gibt es nach Owens eine Standardauffassung. Hinsichtlich der explanatorischen Zustände spiegelt die ‚repräsentationale These‘ die etablierte Meinung wider, bezüglich der erklärungsstiftenden Verbindung erfüllt die ‚kausale These‘ diese Rolle. Die repräsentationale These besagt, dass die Zustände, die Verhalten erklären, propositionale Einstellungen sind bzw. Zustände, die etwas repräsentieren und die ihre Identität über ihre jeweiligen Inhalte beziehen oder deren Unterscheidungsmerkmal in anderen Worten eben ihre Inhalte sind. Die kausale These behauptet, dass psychologische Erklärungen ihrem Wesen nach kausal sind. Das bedeutet, dass Gründe nur dann Verhalten erklären können, wenn das in Frage stehende Verhalten tatsächlich auch durch die genannten Gründe verursacht wurde. Zusammengenommen verkörpern diese beiden Hypothesen das Standarderklärungsmodell der Alltagspsychologie (im Folgenden K/R-Modell), so Owens. Allerdings besteht zwischen diesen beiden Thesen nach Owens ein so starkes Spannungsverhältnis, dass es unmöglich ist beide zugleich kohärent zu vertreten. Die problematische Beziehung der beiden Thesen zueinander wird explizit, wenn man den Gedankenexperimenten von Putnam und Burge folgt. Denn dann ist es nämlich möglich, dass zwei *physikalisch identische* Individuen *Unterschiede in ihren propositionalen Einstellungen* aufweisen. Damit ist aber zugleich die Möglichkeit gegeben, dass psychologische Zustände nicht auf physikalischen Zuständen supervenieren. Um die Unmöglichkeit psychophysischer Supervenienz noch weiter zu untermauern, präsentiert Owens folgendes Argument:

³⁶⁶ Owens, Joseph: „Content, Causation, and Psychophysical Supervenience“ in: *Philosophy of Science*, Vol. 60, 1993, S. 243.

- 1'. Folk psychological explanation makes essential explanatory appeal to propositional attitudes, to states individuated in terms of content (the representational thesis).
- 2'. Folk psychological explanation is in the business of offering causal explanations of behavior (the causal thesis).
- 3'. Explanatory accounts which are in the business of providing causal explanations for the behavior of various systems should distinguish between the explanatory states of such systems only if they differ in causal powers.
- 4'. Hence, folk psychological explanation should distinguish between the psychological explanatory states of the twins, Alf and Alf* (their beliefs, and so on) only if they differ in causal powers.
- 5'. Since the twins are physically identical there is no difference in causal powers between their corresponding explanatory states.
- 6'. Hence, folk psychology should not individuate in such a way as to distinguish between the (ordinary) psychological explanatory states of the twins, that is, it should not individuate in such a way as to justify attributing a particular explanatory state to one twin and not to the other.
- 7'. Hence, belief and desire should not be individuated in such a way as to justify attributing differing beliefs and desires to such twins (by 1' and 6').³⁶⁷

Ausgehend von der repräsentationellen und der kausalen These kommt Owens mit zusätzlichen Prämissen zu der Konklusion, dass die Individuierung intentionaler Zustände physikalisch identischer Zwillinge nicht auf eine Art und Weise erfolgen darf, mittels derer den Zwillingen unterschiedliche intentionale Zustände zugeschrieben werden. Diese Schlussfolgerung kann jedoch von niemanden geteilt werden, der der Kernaussage hinter den Gedankenexperimenten Putnams und Burges beipflichtet. Hier nochmals die einzelnen Schritte von Owens' Argument:

- 1'. Die Erklärungen von Verhalten, die die Alltagspsychologie anbietet, berufen sich in der Regel auf propositionale Einstellungen – also auf Zustände, die über ihre Inhalte individuiert werden (i.e. die repräsentationale These).
- 2'. Die Erklärungen der Alltagspsychologie sind kausale Erklärungen von Verhalten (i.e. die kausale These).
- 3'. Ansätze, die kausale Erklärungen vom Verhalten verschiedener Systeme anbieten, sollten die für die Erklärung genutzten Zustände der in Frage stehenden Systeme nur dann unterscheiden, wenn sich diese in ihren kausalen Kräften unterscheiden.
- 4'. Daher sollten die Erklärungen der Alltagspsychologie die psychologischen Zustände der Zwillinge Alf und Alf* (also ihre Glaubenssätze, etc.), die für die entsprechende Erklärung verwendet werden, nur dann unterscheiden, wenn sich diese Zustände in ihren kausalen Kräften unterscheiden.
- 5'. Da die Zwillinge physikalisch betrachtet identisch sind, können sich die kausalen Kräfte der entsprechenden Zustände, die für die Erklärung genutzt werden, nicht unterscheiden.

³⁶⁷ Owens, Joseph: „Content, Causation, and Psychophysical Supervenience“ in: *Philosophy of Science*, Vol. 60, 1993, S. 248.

- 6'. Folglich sollte die Alltagspsychologie Zustände, die für die kausale Erklärung von Verhalten genutzt werden, nicht so individuieren, dass die psychologischen Zustände der Zwillinge unterschieden werden. Genauer gesagt sollte die Alltagspsychologie nicht auf eine Art und Weise Zustände individuieren, dass damit gerechtfertigt wäre einem der Zwillinge einen explanatorischen Zustand zuzuschreiben, den der andere nicht hat.
- 7'. Schlussendlich sollten Überzeugungen und Verlangen nicht in einer solchen Manier individuiert werden, dass den beiden Zwillingen dadurch unterschiedliche Glaubenssätze und Wünsche zugeschrieben werden können.

Offensichtlich ist dieses Argument problematisch für jeden, der die Intuitionen hinter den Gedankenexperimenten Putnams und Burges teilt. Schließlich ist 7'. das genaue Gegenteil von dem, was vielen intuitiv als wahr erscheint – nämlich, dass der Inhalt von Gedanken, Glaubenssätzen, Wünschen, etc. von zwei physikalisch identischen Zwillingen unter bestimmten Umständen verschieden sein kann. Wer also die von Putnam und Burge entworfenen Szenarien als kohärent empfindet, kann 7'. nicht zustimmen. Da das Argument laut Owens aber gültig ist, muss – wenn 7.' falsch ist – mindestens eine der Prämissen 1'., 2'., 3'. oder 5'. falsch sein.

In seiner Analyse dieser vier Prämissen nimmt Owens zunächst die repräsentationale These (i.e. 1'.) als gegeben hin – daher muss ihm zufolge entweder 2'., 3.' oder 5.' falsch sein. Die Ablehnung von 5.' sei jedoch keine Alternative für einen Verfechter psychophysischer Supervenienz, weil eine Zurückweisung von 5'. dem Grundgedanken hinter jeglicher Vorstellung von Supervenienz widerspräche. Denn unabhängig davon wie Supervenienz genau ausbuchstabiert wird, hat eine Identität auf Ebene der Supervenienzbasis eine Identität auf der Ebene der supervenierenden Eigenschaften zur Folge – genau darin besteht die zentrale Idee jedes Supervenienzkonzepts. Somit bleiben lediglich die Prämissen 2'. und 3'. als Kandidaten für eine Zurückweisung übrig. Owens zufolge ist 3'. intuitiv einleuchtend, denn der dahinterstehende Gedanke besagt, dass ein System, dessen Funktion es ist kausale Erklärungen zu liefern, nicht ein Klassifikationsschema verwenden sollte, das systematisch Zustände unterscheidet, die von einem kausalen Standpunkt aus ununterscheidbar sind. Infolgedessen scheint es als müsse man die These 2'., also die kausale These, fallenlassen.

Abschließend resümiert Owens nochmals den Gedankengang, auf dem sein Argument fußt, so:

To summarize, Alf's token belief that arthritis is crippling is explanatorily linked to his saying that arthritis is crippling, while Alf*'s token belief that tharthrititis is crippling is linked to a different action. This difference in linkage cannot, however, be cashed out in terms of the two tokens being such as to *cause* different behaviors. This is ruled out by (5'), which stipulates that they do not differ in causal powers, and this in turn is forced on us by TPS [Theoretical Psychophysical Supervenience, S.H.]. Apparently, then, proponents of TPS have little or no option but to reject the causal thesis (or, of course, the representationalist thesis).³⁶⁸

Während also Alfs Überzeugung Arthritis rufe Lähmungserscheinungen hervor, eine Erklärung für Alfs Aussage, dass Arthritis Lähmungserscheinungen hervorruft, liefert, verursacht die Überzeugung von Alf* (einem physikalisch identischen Zwilling von Alf), dass Zwarthritis lähmend wirkt, im Gegensatz dazu eine andere Handlung. Allerdings kann an dieser Stelle nicht argumentiert werden, dass der Unterschied, in dem was diese beiden Überzeugungen erklären, darin begründet ist, dass sie unterschiedliche Verhaltensweisen verursachen. Schließlich ist dies durch 5'. ausgeschlossen, wonach sich physikalisch identische Entitäten nicht in ihren kausalen Kräften voneinander unterscheiden können. Dies wiederum ist durch die Annahme der theoretischen psychophysischen Supervenienz (TPS) bedingt. Die Aussage von TPS ist, dass die Zustände, die in der theoretischen Psychologie der Erklärung von Verhalten zugrunde liegen, auf eine Art und Weise individuiert werden müssen, die sicherstellt, dass sich zwei Individuen nur dann in solchen Zuständen unterscheiden können, wenn sie zugleich Unterschiede auf der physikalischen Ebene aufweisen. Somit verbleibt den Anhängern von TPS einzig die Option entweder die kausale oder die repräsentationale These zu verwerfen, so Owens.

Lösungen für das Externalismus/Anti-Individualismus-Problem

Enger Gehalt mentaler Zustände

Im Anschluss an die Gedankenexperimente von Putnam und Burge sieht Jerry Fodor die zeitgenössische Philosophie vor einer Antinomie stehen: so gebe es zwei einander ausschließende Standpunkte, die für sich genommen jedoch beide gleichermaßen einleuchtend sind. In seinem Aufsatz *A Modal Argument for Narrow Content* gibt Fodor diese beiden Sichtweisen in Form zweier Argumente wieder und schildert diesbezüglich auch das zentrale Anliegen seines Essays anhand zweier Leitfragen:

On the one hand, there is argument A:

³⁶⁸ Owens, Joseph: „Content, Causation, and Psychophysical Supervenience“ in: *Philosophy of Science*, Vol. 60, 1993, S. 250f.

Argument A:

1. My twin and I are molecular duplicates.
2. Therefore our (actual and counterfactual) behaviors are identical in relevant respects.
3. Therefore the causal powers of our mental states are identical in relevant respects.
4. Therefore my twin and I belong to the same natural kind for purposes of psychological explanation and “individualism” is true.

But on the other hand, there is argument B:

Argument B

- 1' My twin and I are molecular duplicates.
- 2' Nevertheless, our (actual and counterfactual) behaviors are different in relevant respects.
- 3' Therefore the causal powers of our mental states are different in relevant respects.
- 4' Therefore my twin and I belong to different natural kinds for purposes of psychological explanation and “individualism” is false.

At least one of these arguments must be unsound. Which one? And what is wrong with it?³⁶⁹

Die beiden sich widersprechenden Intuitionen drückt Fodor in Form von zwei Argumenten aus, die er schlichtweg ‚Argument A‘ und ‚Argument B‘ nennt. Dabei weisen beide Argumente dieselbe logische Struktur auf, kommen aber inhaltlich zu diametral entgegengesetzten Schlussfolgerungen. Daher ist nach Fodor mindestens eines der beiden Argumente notwendigerweise nicht schlüssig sein. Dementsprechend besteht die Zielsetzung seines Aufsatzes darin zu klären welches der beiden Argumente fehlerhaft ist (wenn nicht beide gleichermaßen betroffen sind) und worin der entsprechende Makel begründet ist. In seiner Analyse präsentiert Fodor Argument A wie folgt:

1. Mein Zwilling und ich sind molekulare Duplikate.
2. Daher sind unsere (aktualen und kontrafaktischen) Verhaltensweisen im Hinblick auf ihre wesentlichen Belange identisch.
3. Aus diesem Grund sind die kausalen Kräfte unserer mentalen Zustände in ihren entscheidenden Gesichtspunkten identisch.
4. Deshalb gehören mein Zwilling und ich hinsichtlich psychologischer Erklärungen zur selben natürlichen Art und die Theorie des Individualismus ist wahr.

Den Antagonisten zu Argument A – das Argument B – beschreibt Fodor folgendermaßen:

- 1' Mein Zwilling und ich sind molekulare Duplikate.

³⁶⁹ Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 5.

- 2' Gleichwohl unterscheiden sich unsere (aktualen und kontrafaktischen) Verhaltensweisen in wesentlichen Belangen.
- 3' Aus diesem Grund sind die kausalen Kräfte unserer mentalen Zustände in entscheidenden Gesichtspunkten verschieden.
- 4' Deshalb gehören mein Zwilling und ich hinsichtlich psychologischer Erklärungen zu unterschiedlichen natürlichen Arten und die Theorie des Individualismus ist falsch.

Offensichtlich gelangen die Argumente A und B zu kontradiktorischen Konklusionen, obwohl sie von der gleichen Prämisse ausgehen und dieselbe logische Struktur aufweisen. Das oben beschriebene Externalismus-Problem hinsichtlich mentaler Verursachung ergibt sich im Wesentlichen aus einer Interpretation der Gedankenexperimente Putnams und Burges im Sinne von Argument B (oder etwas Vergleichbarem). Im Gegensatz dazu macht sich Fodor jedoch für das Argument A stark, indem er versucht zu beweisen „(...) that mental states that differ only in “broad” intentional properties (the sorts of intentional properties that the mental states of molecular twins may fail to share) ipso facto do not differ in causal powers; hence that mere differences in broad intentional content do not determine differences in natural kinds for purposes of psychological explanation.“³⁷⁰ Allein aus der Tatsache, dass zwei mentale Zustände verschiedene ‚weite‘ intentionale Gehalte besitzen, lasse sich nach Fodor nicht zwangsläufig folgern, dass sich die entsprechenden mentalen Zustände auch in ihren Kausalkräften voneinander unterscheiden. Dementsprechend müssen zwei mentale Zustände, die verschiedene weite intentionale Gehalte besitzen, nicht notwendigerweise zwei unterschiedlichen natürlichen Arten im Hinblick auf psychologische Erklärungen zuzurechnen sein. Offensichtlich besteht die für Fodors Argumentation entscheidende Differenzierung zwischen dem weiten Gehalt und dem engen mentaler Zustände. Während der weite Gehalt mentaler, intentionaler Eigenschaften bei molekular identischen Zwillingen unterschiedlich sein kann, ist dies hinsichtlich ihres engen Gehaltes nicht möglich. Der enge Gehalt intentionaler Eigenschaften ist also an die intrinsischen Eigenschaften der Träger gebunden – sind diese identisch, dann besitzen die intentionalen Eigenschaften auch denselben engen Gehalt. Der enge Gehalt intentionaler Eigenschaften zweier molekular identischer Zwillinge kann dementsprechend auch nicht verschieden sein.

Die grundlegende Frage, die Argument A und Argument B aufwerfen – i.e., ob der Individualismus wahr oder falsch ist – muss Fodor zufolge getrennt von der Frage nach der Wahr- bzw. Falschheit des semantischen Externalismus betrachtet werden. Schließlich stelle sich die Fra-

³⁷⁰ Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 6.

ge nach dem explanatorischen Status des weiten Gehalts intentionaler Eigenschaften unabhängig von der Frage, ob sich der Inhalt meiner Gedanken an Wasser von dem Inhalt der Gedanken meines Zwillings an Zwasser unterscheidet oder nicht. Der zentrale Aspekt, um den die Diskussion über Argument A und B im Wesentlichen oszilliert, ist nach Fodor die Frage, ob der Individualismus wahr oder falsch ist. Dies hängt wiederum von folgender Fragestellung ab: „Do the twins’ mental states belong to different natural kinds (do they have different causal powers) in virtue of the differences in the intentional properties of the twins’ behavior for which they are responsible?“³⁷¹ Es gilt also vornehmlich zu untersuchen, ob die mentalen Zustände der Zwillinge unterschiedlichen oder derselben natürlichen Art angehören oder, ob sie – in anderen Worten – die gleichen oder verschiedene Kausalkräfte besitzen. Die Frage nach dem Besitz von Kausalkräften stellt sich nach Fodor auch in Kontexten, die nichts mit Intentionalität zu tun haben. Ein allgemeines Konzept zur Erfassung für Kausalkräfte könnte Fodor zufolge in etwa so aussehen:

Suppose we have a pair of causes *C1*, *C2*, together with their respective effects *E1*, *E2*. Assume that:

C1 differs from *C2* in that *C1* has a cause property *CP1* where *C2* has cause property *CP2*.

E1 differs from *E2* in that *E1* has effect property *EP1* and *E2* has effect property *EP2*.

The difference between *C1* and *C2* is responsible for the difference between *E1* and *E2* in the sense that, if *C1* had had *CP2* rather than *CP1*, then *E1* would have *EP2* rather than *EP1*; and if *C2* had had *CP1* rather than *CP2*, *E2* would have had *EP1* rather than *EP2*.

Call this *schema S*. And now, what we want to know is: Which instances of *schema S* are cases where the difference between having *CP1* and having *CP2* is a difference in causal power in virtue of its responsibility for the difference between *E1* and *E2*?³⁷²

Sein allgemeines Modell zur Darstellung der Kausalkräfte von Eigenschaften tauft Fodor „Schema S“ und skizziert es wie folgt: Angenommen, man kann den beiden Wirkungen *E1* und *E2* jeweils eine Ursache zuordnen, so dass *C1* die Ursache von *E1* ist und *C2* *E2* verursacht. Zudem gehe man davon aus, dass sich *C1* und *C2* darin unterscheiden, dass *C1* die Eigenschaft *CP1* besitzt, während *C2* die Eigenschaft *CP2* aufweist. Analog gehört die Eigenschaft *EP1* zu Wirkung *E1* und Wirkung *E2* instantiiert Eigenschaft *EP2*. Der Unterschied zwischen *C1* und *C2* sei nunmehr auch für die Verschiedenartigkeit von *E1* und *E2* verantwortlich und zwar in dem Sinn, dass, wenn *C1* anstelle der Eigenschaft *CP1* die Eigenschaft *CP2* besessen hätte, *EP2* (anstelle von *EP1*) die zu *E1* gehörige Eigenschaft gewesen wäre. Gleichermäßen hätte *E2* die Eigenschaft *EP1* (und nicht *EP2*) instantiiert, wenn *C2* die Eigenschaft *CP1* anstelle von *CP2* besessen hätte. Für eine adäquate Untersuchung vor dem Hintergrund von Argument A und Argument B gilt es nach Fodor zu klären welche Vorkommnisse

³⁷¹ Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 9.

³⁷² Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 9.

von Schema S so beschaffen sind, dass der Unterschied zwischen dem Haben der Eigenschaft CP1 und dem Besitzen der Eigenschaft CP2 einen Unterschied der kausalen Kräfte im Hinblick auf den Unterschied zwischen E1 und E2 zur Folge hat. Denn nicht in jedem Fall hat der Besitz unterschiedlicher Eigenschaften auch notwendigerweise das Haben verschiedener Kausalkräfte zur Folge. Dazu gibt Fodor folgendes, triviales Beispiel: „One of the properties of my effects that your effects cannot have, however hard you try, is *the property of being caused by me*. But I take it to be simply obvious that this difference in our effects does not make the property of *being me rather than you* a causal power. I am not a unit natural kind in virtue of my unique power to cause effects that are effects of me.“³⁷³ Eine Eigenschaft, die von mir verursachte Wirkungen haben, die Wirkungen, die von jemand anderem hervorgebracht werden, keinesfalls besitzen können ist die Eigenschaft von mir verursacht worden zu sein. Dies macht jedoch gewiss nicht die Eigenschaft ich anstatt jemand anderer zu sein zu einer kausalen Kraft. Schließlich bin ich keine einheitliche, natürliche Art dank meines einzigartigen Vermögens Wirkungen zu verursachen, die meine Wirkungen sind. Daraus zieht Fodor dieses Fazit: „It seems clear a priori, then, that not every case in which a difference between causes is responsible for a difference in effects is a case where the difference in causes is a difference in their causal powers.“³⁷⁴ Demnach ist es für Fodor offensichtlich, dass nicht jedes Mal, wenn ein Unterschied zwischen den Ursachen für einen Unterschied bei den Wirkungen verantwortlich ist, der Unterschied zwischen den Ursachen einem Unterschied derer Kausalkräfte gleichzusetzen ist. Ein weiteres Exempel, das Fodor in diesem Zusammenhang zur Untermauerung seines Gedanken darbietet ist dieses: kraft meines Habens von Geschwistern, ist es mir möglich Söhne zu haben, die Neffen sind. Ein mit mir auf molekularer Ebene identischer Zwilling, der Einzelkind ist, könnte dementsprechend keine männlichen Nachkommen haben, die zugleich Neffen väterlicherseits sind. Allerdings scheint es evident zu sein, dass die Eigenschaft Geschwister zu haben bzw. die Eigenschaft keine Geschwister zu haben nicht einen Unterschied in den Kausalkräften von Eltern darstellt vermöge ihrer Verantwortlichkeit für den Unterschied in den Eigenschaften ihrer Zöglinge. An dieser Stelle könnte man versucht sein zu schließen, dass der Unterschied zwischen CP1 und CP2 in diesem Vorkommnis von Schema S keinen Unterschied in den Kausalkräften repräsentiert, weil CP1 und CP2 in diesem Fall relationale Eigenschaften sind, so Fodor. Allerdings zeigt sich nach kurzer Überlegung, dass dies nicht der Fall sein kann. In der Wissenschaft ist der Gebrauch relationaler Eigenschaften schließlich weit verbreitet und es steht außer Frage, dass solche Eigenschaften (manchmal) auch kausale Kraft besitzen. Betrachtet man z.B. die Eigen-

³⁷³ Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 11.

³⁷⁴ Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 11.

schaft ein Planet zu sein und die Eigenschaft ein Meteor zu sein, so sind dies eindeutig relationale Eigenschaften, denn es könnten zwei molekular identische Gesteinsbrocken existieren, von denen der eine ein Planet und der andere ein Meteor ist, je nachdem in welche Umwelt der jeweilige Gesteinsklumpen eingebettet ist. Unter diesen Umständen könnte man etwa wahrheitsgemäß Folgendes behaupten: da dieser Gesteinsbrocken ein Planet ist und der andere Gesteinsbrocken kein Planet ist, befindet sich ersterer auf einer keplerschen Umlaufbahn und letzterer nicht. Andererseits hat jener Gesteinsbrocken, weil er ein Meteor ist, unter seinen (potentiellen) Wirkungen Krater, während der andere Gesteinsbrocken – also der Planet – nicht die kausale Kraft besitzt in der Lage zu sein Krater zu verursachen.³⁷⁵ Als Quintessenz aus diesen Beispielen konstatiert Fodor zusammenfassend: „(...) not all cause properties are causal powers; not all relational properties fail to be (...).“³⁷⁶ Demgemäß besitzen nicht alle Eigenschaften einer Ursache kausale Kraft und bisweilen haben auch relationale Eigenschaften Kausalkraft. Um eine Möglichkeit zur Identifikation der aktuellen kausalen Kräfte aufzuzeigen, verweist Fodor auf folgende Eigentümlichkeit von Schema S:

(...) the question we are raising is not whether the difference between having *CP1* and having *CP2* is a difference in causal powers; rather, it is whether the difference between having *CP1* and having *CP2* is a difference in causal powers *in virtue of its being responsible for a certain difference between *E1* and *E2**, namely, in virtue of its being responsible for *E1*'s having *EP1* rather than *EP2* and for *E2*'s having *EP2* rather than *EP1*. The point I am wanting to emphasize is that a cause property might fail to count as a causal power in virtue of its responsibility for one effect property, but still might constitute a causal power in virtue of its responsibility for some other effect property.³⁷⁷

Somit ist es nicht entscheidend, ob der Unterschied zwischen dem Haben von *CP1* und dem Besitz von *CP2* in verschiedenen kausalen Kräften resultiert, sondern es ist vielmehr von Bedeutung, ob dieser Unterschied andersgeartete Kausalkräfte aufgrund seiner Verantwortlichkeit für einen bestimmten Unterschied zwischen *E1* und *E2* bedingt. Genauer gesagt ist es entscheidend, ob *E1* die Eigenschaft *EP1* und nicht *EP2* besitzt, respektive ob *E2* die Eigenschaft *EP2* anstelle von *EP1* aufweist, weil *CP1* und *CP2* verschiedene Kausalkräfte besitzen. Allgemein gesprochen ist es also durchaus möglich, dass eine beliebige Eigenschaft einer Ursache im Hinblick auf ihre Verantwortlichkeit für eine bestimmte Wirkungseigenschaft keine kausale Kraft besitzt, während dieselbe Ursacheneigenschaft hingegen durchaus kausale Kraft hinsichtlich einer anderen Eigenschaft der Wirkung haben kann. So ist beispielsweise die Eigenschaft Geschwister zu haben keine Kausalkraft in dem Sinne, dass sie mir es ermöglicht Söhne zu haben, die Neffen sind. Aber angenommen, es gäbe eine ‚Geschwister-

³⁷⁵ vgl. Fodor, Jerry: „*A Modal Argument for Narrow Content*“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 11f.

³⁷⁶ Fodor, Jerry: „*A Modal Argument for Narrow Content*“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 12.

³⁷⁷ Fodor, Jerry: „*A Modal Argument for Narrow Content*“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 12f.

Krankheit', die bei denjenigen Menschen, die Geschwister haben einen Ausschlag hervorruft, dann könnte die Eigenschaft Geschwister zu besitzen durchaus eine kausale Kraft darstellen. Laut Fodor gilt es bei der Analyse von Kausalkräften grundsätzlich zwei Kategorien von Fällen zu unterscheiden. Die erste Klasse umfasst die Fälle, in denen das Haben der kausalen Kraft nicht-kontingent ist, die zweite diejenigen, in denen es kontingent ist. D.h. in den Szenarien des ersten Typs, ist CP1 beispielsweise selbst die Eigenschaft die kausale Kraft zu besitzen Ereignisse verursachen zu können, die EP1 haben. Offensichtlich ist unter diesen Umständen der Unterschied zwischen dem Haben und Nichthaben von CP1 dem Unterschied eine bestimmte Kausalkraft aufzuweisen oder nicht gleichzusetzen. Ein ähnlich gelagerter Fall wäre etwa die Eigenschaft wasserlöslich zu sein – auch in diesem Fall ist der Besitz kausaler Kraft bereits durch die Beschaffenheit der Eigenschaft gegeben, oder in anderen Worten nicht-kontingent. Denn alle Dinge, die diese Eigenschaft aufweisen, haben die kausale Kraft sich in Wasser aufzulösen. Bei den Eigenschaften ein Planet oder ein Meteor zu sein ist das Haben von Kausalkraft hingegen kontingent. D.h. diese Eigenschaften besitzen kausale Kräfte aufgrund ihres Vermögens keplersche Umlaufbahnen bzw. Krater hervorzurufen. Da es jedoch kontingent ist, dass Planeten auf keplerschen Umlaufbahnen kreisen und dass Meteore Krater bewirken, kann man auch den Besitz der Kausalkräfte dieser Eigenschaften als kontingent betrachten.³⁷⁸ In Anwendung auf die Individualismusdebatte zieht Fodor folgende Parallele:

Notice that the broad content cases are like the meteor and planet cases and unlike the dispositional and functional cases. It may be that being connected to water rather than twater (hence having water thoughts rather than twater thoughts) is having a causal power; but if it is, it is contingent that it is. The property of being connected to water is not *identical* to the property of having a certain causal power, though it may be that there are causal powers that one has if one is connected to water that one would not have if one were not.³⁷⁹

Der weite Gehalt intentionaler, mentaler Eigenschaften ist nach Fodor nicht mit funktionalen oder dispositionalen Eigenschaften, sondern vielmehr mit Eigenschaften wie der ein Planet oder ein Meteor zu sein, vergleichbar. Denn selbst wenn es unterschiedliche Kausalkräfte nach sich zieht mit Wasser anstatt mit Zwasser in Kontakt zu sein – d.h. Wasser-Gedanken anstelle von Zwasser-Gedanken zu haben – so ist dies kontingenter Natur. Denn die Eigenschaft mit Wasser verbunden zu sein ist nicht identisch mit der Eigenschaft eine bestimmte kausale Kraft zu besitzen.

Daraus zieht Fodor zwei Konklusionen: Erstens hat die Tatsache, dass das Haben einer funktionalen Eigenschaft mit dem Besitz einer bestimmten Kausalkraft gleichzusetzen ist, keiner-

³⁷⁸ vgl. Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 15f.

³⁷⁹ Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 16.

lei Auswirkungen auf die Frage, ob auch das Haben einer Eigenschaft mit weitem Gehalt notwendigerweise den Besitz einer Kausalkraft nach sich zieht. Diese Schlussfolgerung mag zunächst paradox erscheinen, denn falls psychologische Eigenschaften funktionale Eigenschaften sind (wovon viele Philosophen ausgehen) und funktionale Eigenschaften nicht-kontingenterweise eine spezifische kausale Kraft zur Folge haben, dann stellt sich die Frage wie es sein kann, dass das Haben Wasser-Gedanken bzw. Zwasser-Gedanken keine kausale Kraft sind. Dies bestreitet Fodor nicht – ihm zufolge impliziert das Haben von Wasser-Gedanken unbedingt eine Kausalkraft (z.B. könnten die Wasser-Gedanken kausal verantwortlich für mein Greifen nach einem mit Wasser gefüllten Gefäß sein) genauso wie das Haben von Zwasser-Gedanken Kausalkraft innehaben kann (die Zwasser-Gedanken meines Zwilings könnten etwa kausal verantwortlich für seine Handbewegung zum Zwasser-Glas sein). Die Frage, die für lokale Supervenienz und folglich auch für der Individualismus-Debatte essentiell ist, lautet jedoch, ob ein Unterschied zwischen Wasser- und Zwasser-Gedanken einen Unterschied in deren Kausalkraft bewirkt. Ein Anti-Individualist bezieht in dieser Angelegenheit einen affirmativen Standpunkt – Wasser- und Zwasser-Gedanken haben aufgrund des intentionalen (beschreibbaren) Unterschieds zwischen den Verhaltensweisen, die sie jeweils verursachen, verschiedene Kausalkräfte. Ein Individualist hingegen negiert dies mit dem Verweis darauf, dass Wasser- und Zwasser-Gedanken dieselbe kausale Kraft besitzen, aber jeweils in Personen mit verschiedenen kausalen Geschichten realisiert wurden. Die zweite Konsequenz, die Fodor aus den oben genannten Punkten zieht, besteht in Behauptung, dass man sich bei der Suche nach einer brauchbaren Bedingung zur Identifikation der kausalen Eigenschaften, die einen Unterschied in den Kausalkräften nach sich ziehen, auf die Fälle beschränken kann, in denen es kontingent ist, ob der Unterschied zwischen den Eigenschaften einen Unterschied zwischen den kausalen Kräften darstellt.³⁸⁰

Aufbauend auf diese Erkenntnisse erklärt Fodor warum beispielsweise die Eigenschaft ein Planet zu sein eine kausale Kraft ist während etwa die Eigenschaft Geschwister zu haben keine solche ist.

So, then: Here is me and here is my molecular twin; and I have siblings and he does not; and in virtue of my having siblings my sons are nephews, and in virtue of his not having siblings his sons are not nephews; and what we want to know is: Why is *having siblings* not a causal power in virtue of its being responsible for this difference in our offspring? Here is a first fling at the answer: it is because having siblings is *conceptually* connected to having sons who are nephews; to be a nephew just is to be a son whose parents have siblings. And, to put it roughly, your causal powers are a function of your *contingent* connections, not of your conceptual connections. (...)

Compare cases of relational properties that really are causal powers, like *being a planet*. Being a planet is a causal power in virtue of, for example, its contingent (a fortiori, nonconceptual) connection with having a Keplerian

³⁸⁰ vgl. Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 16ff.

orbit. That is, being a planet is a causal power because it is true and contingent that, if you have molecularly identical chunks of rock, one of which is a planet and the other of which is not, then, *ceteris paribus*, the one which is a planet will have a Keplerian orbit, and *ceteris paribus*, the one which is not a planet will not.³⁸¹

Die Antwort auf die Frage warum die Eigenschaft Geschwister zu haben im Gegensatz zur Eigenschaft ein Planet zu sein keine Kausalkraft repräsentiert, liegt laut Fodor in der Art und Weise wie die Eigenschaft mit der Wirkung verknüpft ist. Die Eigenschaft Geschwister zu haben ist demnach keine kausale Kraft, weil das Haben von Geschwistern *begrifflich* mit dem Haben von Söhnen, die Neffen sind, verbunden ist. In anderen Worten bedeutet ein Neffe zu sein nämlich nicht anderes, als ein Sohn zu sein, von dem (mindestens) ein Elternteil einen Bruder oder eine Schwester hat. Allgemein gesprochen hängen aber kausale Kräfte nicht von den begrifflichen, sondern vielmehr von den kontingenten Beziehungen ab. Aus diesem Grund ist z.B. die Eigenschaft ein Planet zu sein, im Gegensatz zur Eigenschaft Geschwister zu haben, eine kausale Kraft. Da nämlich die Verbindung mit dem Haben einer keplerschen Umlaufbahn kontingenter Natur ist – und somit nicht schon allein begrifflich bedingt ist – besitzt die Eigenschaft ein Planet zu sein Kausalkraft. Anders ausgedrückt ist die Eigenschaft ein Planet zu sein deshalb eine Kausalkraft, weil es wahr und kontingent ist, dass, wenn von zwei molekular identischen Gesteinsbrocken, der eine ein Planet ist und der andere kein Planet ist, folgendes gilt: der eine Gesteinsklumpen, der ein Planet ist, besitzt *ceteris paribus* eine keplersche Umlaufbahn, wohingegen der andere Gesteinsbrocken unter sonst gleichen Bedingungen keine solche besitzt.

Ausgehend von diesen Beispielen abstrahiert Fodor folgende allgemein gültige Formel als Voraussetzung für das Haben kausaler Kraft:

Here is the general form of the proposed solution. Consider an instance of schema *S*. *C1* has *CP1*, *C2* has *CP2*, *E1* has *EP1*, *E2* has *EP2*, and the difference between the causes is responsible for the difference between the effects in the sense that *E1* would not have had *EP1* (rather than *EP2*) but that *C1* had *CP1* (rather than *CP2*). And what we want to know is: When does the fact that this difference in the causes is responsible for this difference in the effects make *CP1* and *CP2* causal powers? The answer, which I shall call *condition C*, is:

Only when it is not a conceptual truth that causes which differ in that one has *CP1* where the other has *CP2* have effects that differ I that one has *EP1* where the other has *EP2*.³⁸²

Im Schema *S* sind *CP1* und *CP2* also genau dann als Kausalkräfte einzustufen, wenn der Unterschied zwischen den Eigenschaften *CP1* und *CP2* der Ursachen *C1* und *C2* für den Unterschied zwischen den Eigenschaften *EP1* und *EP2* der Wirkungen *E1* und *E2* verantwortlich ist und darüber hinaus auch die Bedingung *C* erfüllt ist. Die Bedingung *C* besagt, es darf keine rein begriffliche Wahrheit sein, dass Ursachen, die sich in dem Punkt unterscheiden, dass die

³⁸¹ Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 19.

³⁸² Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 19.

eine Ursache Eigenschaft CP1 aufweist während die andere die Eigenschaft CP2 besitzt, Wirkungen verursachen, von denen die eine die Eigenschaft EP1 aufweist, wohingegen die andere EP2 hat. Genauer gesagt erachtet Fodor die Bedingung C als notwendige, aber nicht als hinreichende, Bedingung für den Besitz kausaler Kraft. Im Umkehrschluss gilt folglich, dass, wenn eine Eigenschaft im Schema S die Bedingung C nicht erfüllt, diese Eigenschaft demzufolge auch keine kausale Kraft (im Hinblick auf eine bestimmte andere Eigenschaft) besitzt. Laut Fodor erfüllen Eigenschaften mit weitem Gehalt die Bedingung C nicht und dies genüge um den Individualismus zu rechtfertigen. Dazu argumentiert Fodor wie folgt: Betrachtet man meine Eigenschaft Wasser in meiner Geschichte zu haben – i.e. die Eigenschaft auf die Art und Weise mit Wasser verbunden zu sein wie ich es bin und wie es mein molekularer Zwilling nicht ist. Der Unterschied zwischen dieser Art von Verbindung und dem Fehlen dieser Verbindung ist verantwortlich für einen bestimmten Unterschied zwischen dem weiten Gehalt meiner Gedanken und dem weiten Gehalt der Gedanken meines Zwillings – genauer gesagt hat dieser Unterschied zur Folge, dass ich Wasser-Gedanken habe, während mein Zwilling Zwasser-Gedanken hat. Nun stellt sich die Frage, ob der Unterschied zwischen meiner Geschichte und der meines Zwillings eine Kausalkraft darstellt aufgrund des Unterschieds zwischen den Gehalten unserer Gedanken, für die er verantwortlich ist. Gemäß Fodor ist dies nicht der Fall, denn es ist *begrifflich notwendig*, dass wenn man in der richtigen Art von Beziehung zu Wasser steht, man Wasser-Gedanken anstatt Zwasser-Gedanken hat (und umgekehrt). Denn im Grunde ist das Haben von Wasser-Gedanken gleichbedeutend damit einen Gedanken zu haben, der in einer richtigen Beziehung zu Wasser steht; analog ist das Haben von Zwasser-Gedanken eigentlich nichts anderes als einen Gedanken zu denken, der in einer passenden Relation zu Zwasser steht. Dementsprechend ist es nicht der Fall, dass mein Verbundensein mit Wasser anstelle mit Zwasser einen Unterschied in meinen kausalen Kräften aufgrund meines Habens von Wasser-Gedanken anstatt von Zwasser-Gedanken zur Folge hat. Obwohl das Haben von Wasser-Gedanken Wasser-Verhalten (wie z.B. das Graben nach Wasser) verursacht und das Denken von Zwasser-Gedanken demgegenüber Zwasser-Verhalten (beispielsweise das Graben nach Zwasser) hervorruft, zählt der Unterschied zwischen beiden nicht als kausale Kraft aufgrund der Tatsache, dass er für den Unterschied bei den intentionalen Eigenschaften des Verhaltens der jeweiligen Träger der Gedanken verantwortlich ist. Denn es ist vielmehr begrifflich notwendig, dass jemand, der Wasser-Gedanken (und keine Zwasser-Gedanken) hat Wasser-Verhalten (anstelle von Zwasser-Verhalten) an den Tag legt. Schließlich ist Wasser-Verhalten im Gegensatz zu Zwasser-Verhalten dadurch gekennzeichnet, dass es von Wasser-Gedanken und nicht von Zwasser-Gedanken verursacht wurde. So ist

es zwar wahr, dass Wasser-Gedanken Wasser-Verhalten erzeugen und Zwasser-Gedanken dies nicht tun, daraus folgt aber nicht, dass Wasser-Gedanken eine kausale Kraft besitzen die Zwasser-Gedanken fehlt.³⁸³ Vielmehr gilt laut Fodor: „(...) being a water thinker is the same causal power as being a twater thinker, only instantiated in a person with a different causal history.“³⁸⁴ Kurzum, Wasser- und Zwasser-Gedanken besitzen dieselben Kausalkräfte und sind lediglich das Produkt von unterschiedlichen kausalen Geschichten der Denker der jeweiligen Gedanken.

Zusammenfassend resümiert Fodor seinen Argumentationsgang folgendermaßen:

So, here is the story. For the difference between being *CP1* and being *CP2* to be a difference in causal powers, it must at least be that the effects of *CP1* differ from the effects of being *CP2*. But, I claim, it is further required that this difference between the effects be *nonconceptually* related to the difference between the causes. This further condition is motivated both by our intuitions about the examples and by Humean consideration that causal powers are, after all, powers to enter into nonconceptual relations. Broad-content differences, per se, do not satisfy this condition, however. There are differences between my behavior and my twin's which are due, in the first instance, to the difference between the intentional contents of our thoughts, and, in the second instance, to my being connected to water in a way that he is not. But these differences among the effects are conceptually related to the differences between the causes; it is conceptually necessary that being connected to water rather than twater leads to water thinking rather than twater thinking; and it is again conceptually necessary that water thinking leads to water behaving and twater thinking does not.

So, then, the difference between the mental states of the twins does not count as a difference in causal power in virtue of its responsibility for the intentional differences among twin behaviors. So argument *B* is no good; what is wrong with it is that the inference from 2' to 3' is unsound. Finally, since it is assumed that the effects of mental states that differ only in broad content are (relevantly) different *only* under intentional description, it follows that there are no taxonomically relevant differences consequent upon broad-content differences as such. From the point of view of psychological taxonomy, my mental states must therefore belong to the same natural kind as those of my molecular twin. So individualism is true and local supervenience is preserved. End of story.³⁸⁵

Ein Unterschied in den Kausalkräften von CP1 und CP2 erfordert erstens, dass sich die Wirkungen von CP1 von denen von CP2 unterscheiden und zweitens, dass der Unterschied zwischen den Wirkungen nicht begrifflich mit dem Unterschied zwischen den Ursachen verknüpft ist, so Fodor. Allerdings werden Unterschiede im weiten Gehalt von Eigenschaften dieser zweiten Bedingung nicht gerecht. Die Unterschiede im Verhalten meines Zwillings und in meinem Verhalten sind einerseits den intentionalen Gehalten unserer Gedanken zu verdanken, andererseits aber auch dem Umstand, dass ich in einer besonderen Art und Weise mit Wasser in Verbindung stehe, in der mein Zwilling nicht steht. Allerdings besteht ein begrifflicher Zusammenhang zwischen den Unterschieden in den Wirkungen und den Unterschieden in den Ursachen. Denn es ist begrifflich notwendig, dass man Wasser-Gedanken (und keine Zwasser-Gedanken) hat, wenn man mit Wasser statt mit Zwasser im Kontakt steht. Analog führt auch das Haben von Wasser-Gedanken zu Wasser-Verhalten, was man von Zwasser-

³⁸³ vgl. Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 20f.

³⁸⁴ Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 20.

³⁸⁵ Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 24f.

Gedanken nicht behaupten kann. Daher kann der Unterschied zwischen den mentalen Zuständen der Zwillinge nicht als ein Unterschied in den daraus resultierenden Kausalkräfte gewertet werden. Aus diesem Grund ist auch das anfangs beschriebene Argument B fehlerhaft. Genauer gesagt ist der Schluss von 2' auf 3' (also von den Unterschieden in den aktuellen und kontrafaktischen Verhaltensweisen der Zwillinge auf Unterschiede in den Kausalkräften ihrer mentalen Zustände) laut Fodor nicht zulässig. Da zudem angenommen wurde, dass sich die Wirkungen mentaler Zustände, die sich nur im weiten Gehalt unterscheiden, ausschließlich unter intentionaler Beschreibung der Zustände verschieden sind, folgt, dass es keine wichtigen taxonomischen Unterschiede gibt, die aus den Unterschieden zwischen den weiten Gehalten resultieren. Für das psychologische Begriffsinventar bedeutet das, die mentalen Zustände meines Zwillings und mir gehören derselben natürlichen Art an. Somit ist auch der Individualismus wahr und gleichzeitig die lokale Supervenienz gewahrt.

Der letzte Schachzug, den Fodor in seinem Essay vornimmt, ist die Umwandlung dieses Gedankengangs in ein Argument für den engen Gehalt intentionaler Zustände und dies bewerkstelligt er so:

We have seen that twater thoughts and water thoughts are not different causal powers. So, for the psychologist's purposes, they are the same intentional state. But they cannot be the same intentional state unless they have the same intentional content. And they cannot have the same intentional content unless intentional content is individuated narrowly. Now it is an argument for narrow content.³⁸⁶

Da Wasser- und Zwasser-Gedanken keine unterschiedlichen Kausalkräfte besitzen, sind sie vom psychologischen Standpunkt aus betrachtet ein und derselbe intentionale Zustand. Allerdings können sie nicht ein und derselbe intentionale Zustand sein, sofern sie nicht denselben intentionalen Gehalt haben. Dies ist jedoch nur dann möglich, wenn der intentionale Gehalt mentaler Zustände eng individuiert wird.

Erklärung vs. Verursachung

Wie oben bereits ausführlich dargelegt nutzt Dretske eine Analogie zwischen dem Verhalten von Münzautomaten und menschlichen Verhalten, um das von Externalismus und Anti-Individualismus aufgeworfene Problem für mentale Verursachung darzustellen. Hierzu nochmals Dretskes Beschreibung der Folgen einer Akzeptanz dieser Analogie, ergänzt um seinen persönlichen Standpunkt hinsichtlich dieser Problematik:

³⁸⁶ Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991, S. 25.

If we take this analogy seriously, it suggests that beliefs, qua beliefs, are as irrelevant to animal behavior as is money, qua money, to the behavior of vending machines. Since it is facts about the shape and size of coins, not facts about their monetary value, that explain why coins cause a machine to yield its contents, the analogy, if we take it seriously – and a good many philosophers do – compels us to conclude that it is the intrinsic features of belief, their neurobiological properties, not their extrinsic properties, their meaning or content, that explains why we do what we do. We thus seem to be driven to the conclusion that what we believe is causally irrelevant to what we do.

I do not think we are *driven* to this conclusion, although, I admit, some people seem willing to drive there. It is the purpose of this essay to say why this conclusion is not forced on us.³⁸⁷

Beherzigt man also die Kernaussage dieser Analogie, dann sind ausschließlich intrinsische Eigenschaften für kausale Vorgänge verantwortlich. Da aber z.B. der Inhalt und die Bedeutung von Glaubenssätzen extrinisch individuiert werden, spielen sie unter diesen Rahmenbedingungen letztendlich keinerlei Rolle für die Erklärung von Handlungen. Demnach scheint es als wäre man zur Schlussfolgerung verurteilt, dass das, was man glaubt, kausal irrelevant ist für das, was man tut. Dretske sieht sich jedoch nicht verpflichtet dieser Konklusion zuzustimmen.

Um diesen Gedanken auszuarbeiten, weitet Dretske zunächst seine Analogie aus. So ist beispielsweise die Regierung der Vereinigten Staaten von Amerika sehr darum bemüht, dass der Zahlungswert von Objekten auf den intrinsischen Eigenschaften der Objekte superveniert, die diesen Wert besitzen. Um Fälschungen vorzubeugen werden Banknoten mit diversen Sicherheitsmerkmalen versehen – mitunter Wasserzeichen, Spezialpapier, Sicherheitsfaden, Tiefdruckverfahren, Mikroschrift, usw. Aus diesem Grund gibt es nahezu keine Fälschungen von \$20-Noten, die sich tatsächlich so anfühlen und die so aussehen wie ein echter \$20-Schein. Allerdings ist der Geldwert eines Zahlungsmittels bekanntermaßen keine intrinsische Eigenschaft, sondern hat vielmehr mit dessen Geschichte und mit gesellschaftlichen Konventionen zu tun. Genauer gesagt, hängt der Geldwert im Wesentlichen von der Herkunft des Geldes ab. Die entscheidende Fragestellung in diesem Zusammenhang lautet, wo eine Banknote gedruckt bzw. eine Münze geprägt wurde – in einer staatlichen Prägeanstalt oder in einer illegalen Fälscherwerkstätte. Wenn die Regierung gute Arbeit leistet, dann gilt jedoch: was aussieht und sich anfühlt wie eine \$20-Note ist tatsächlich eine 20\$-Note.³⁸⁸ Eine andere Art diesen Umstand zu beschreiben ist nach Dretske: „(...) that (as long as counterfeiting is kept in check) monetary values (V) are necessarily realized (usually multiply realized) in an object's intrinsic properties (S), and each value of S has the same value of V. This corresponds to what Kim

³⁸⁷ Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 259f.

³⁸⁸ vgl. Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 260.

calls *weak supervenience*.“³⁸⁹ Allgemein gilt, wenn das Fälschen von Geld weitestgehend unter Kontrolle ist (d.h. wenn sich kaum nennenswerte Mengen von schwer identifizierbaren Falschgeld in Umlauf befinden), dann wird der Geldwert V eines Objekts notwendigerweise durch dessen intrinsische Eigenschaften S verwirklicht, so dass jeder Wert von S denselben Wert V besitzt. Dies wiederum entspricht Kims Definition von schwacher Supervenienz. Aufgrund dieser Supervenienz und den damit einhergehenden Korrelationen, wird die Tatsache, dass etwas Geld ist, oft dazu genutzt um die Wirkungen von Geld in verschiedenen Arten von Geschäftsabwicklungen und Transaktionen vorherzusagen und zu erklären. So könnte beispielsweise die Frage, warum mir die Kassiererin \$8 Wechselgeld herausgegeben hat, beantwortet werden, indem man sagt: Weil das Mittagessen \$12 gekostet hat und ich ihr einen \$20-Schein gegeben habe. Obwohl Erklärungen wie diese sicherlich den Alltagsgebrauch widerspiegeln, stellt sich nichtsdestotrotz die Frage, ob sie korrekt sind. Ist die Tatsache, dass ich der Verkäuferin \$20 überreicht habe, wirklich eine Erklärung dafür, dass sie mir \$8 Restgeld herausgegeben hat? In anderen Worten, ist der Geldwert des Papierstücks, das ich ihr ausgehändigt habe, de facto eine kausal relevante Eigenschaft?³⁹⁰ In diesem Kontext macht Dretske auf eine wichtige Unterscheidung aufmerksam:

It is important to understand that these are questions about the causal relevance of an object's properties (its being worth \$20), not the causal efficacy of the objects (the \$20 bills) that have these properties. These are, in other words, questions about what *explains* the result, not what *causes* it. Giving the cashier an object with a monetary value of \$20 *caused* her to give me \$8 change. About that there is no argument. The question we are asking, though, is not whether a \$20 bill is a causally effective *object*, but whether its being a \$20 bill explains its effectiveness. Is the value of the paper I give her a fact about the paper that explains the result of giving her the paper? What if I, instead, give her a piece of paper that looks and feels exactly like a real \$20 bill? Would the result be different if we suppose the bill was a perfect counterfeit? No, of course not. If she can't tell the difference, how could it be? Well, if we really believe this, as I assume we all do, then why say that the cashier gave me \$8 change *because* I gave her \$20? Giving her \$20 is the cause, but that it was \$20 is not the explanation of her giving me \$8 change. The correct explanation is that I gave her a piece of paper that looked and felt (to her) like a \$20 bill. The causally effective properties, those that explain why the effect occurs, are the intrinsic, the observable properties of the paper on which its being \$20 supervenes, the properties you and I, cashiers and machines, use to *tell* whether it is \$20.³⁹¹

Die oben genannten Fragen beziehen sich also auf die kausale Relevanz einer Eigenschaft des Objekts (i.e. \$20 wert zu sein) und nicht auf die kausale Wirksamkeit des Objekts selbst (die \$20-Note), das diese Eigenschaft aufweist. Anders ausgedrückt geht es bei diesen Fragen darum in Erfahrung zu bringen, was die Wirkung *erklärt* und nicht darum aufzuzeigen, was sie *verursacht*. Denn es besteht nämlich kein Zweifel daran, dass mein Bezahlen mit einem \$20-

³⁸⁹ Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 260f.

³⁹⁰ vgl. Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 261.

³⁹¹ Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 261.

Schein die Ursache dafür war, dass mir die Dame an der Kasse \$8 Wechselgeld herausgegeben hat. Die Frage war aber nicht, ob der \$20-Schein das kausal wirksame Objekt war, sondern vielmehr, ob seine Eigenschaft \$20 wert zu sein seine kausale Wirksamkeit erklärt. Allerdings ist anzunehmen, dass die Kassiererin mir auch dann \$8 Restgeld ausgehändigt hätte, wenn ich ihre eine perfekte Fälschung eines \$20-Scheins gegeben hätte. Da die Blüte per definitionem von einem Laien nicht von einer echten Banknote zu unterscheiden gewesen wäre, hätte die Kassiererin entsprechend auch keinen solchen Unterschied entdecken können. Wenn dies der Fall ist, dann stellt sich jedoch die Frage, warum im Alltag Erklärungen wie ‚Die Kassiererin hat mir \$8 Wechselgeld herausgegeben, *weil* ich ihr \$20 überreicht habe‘ so weit verbreitet sind. Denn mein Bezahlen mit \$20 ist zwar gewiss die Ursache dafür, dass sie mir \$8 aushändigt, aber die Tatsache, dass es \$20 waren, ist nicht gleichzeitig auch eine Erklärung hierfür. Die zutreffende Erklärung wäre vielmehr, dass ich ihr ein Stück Papier überreicht habe, das für die Kassiererin wie eine \$20-Note ausgesehen und sich auch so angefühlt hat. Die kausal wirksamen Eigenschaften sind also die beobachtbaren und intrinsischen Eigenschaften des Stücks Papier, auf denen seine Eigenschaft \$20 wert zu sein superveniert. Diese Eigenschaften werden gemeinhin auch dazu genutzt festzustellen, ob es sich bei einem Schein um eine echte \$20-Note handelt, oder um eine \$20-Blüte.

Nun stellt sich allerdings die Frage warum die explanatorische Praxis so ist, wie sie ist – d.h. weshalb werden häufig extrinsische (kausal nicht wirksame) Eigenschaften bei Erklärungen intrinsischen (kausal wirksamen) Eigenschaften vorgezogen? Dretske argumentiert in diesem Zusammenhang, dass es gar nicht nötig sei die explanatorische Praxis zu ändern, denn meistens wären wir uns bewusst, dass die extrinsischen, im Rahmen der Erklärung genutzten, Eigenschaften das Ergebnis im eigentlichen Sinne gar nicht erklären. Wenn man, um beispielsweise das Verhalten eines Verkaufsautomaten vorauszusagen, den Wert von Münzen erwähnt (‚Für eine Cola musst Du \$0,75 einwerfen‘), dann weiß man in der Regel, dass nicht der Wert der Münzen den Auswurf des Getränks erklärt, sondern vielmehr die intrinsischen Eigenschaften der Geldstücke wie Dichte, Größe oder Gewicht. Verallgemeinert gesprochen, würde ein Objekt mit denselben S, aber unterschiedlichem V (in anderen Worten eine gute Fälschung) dasselbe Verhalten hervorrufen. Zudem sind Verkaufsautomaten so aufgebaut, dass sie die intrinsischen Eigenschaften der eingeworfenen Metallstücke überprüfen und nicht deren Entstehungsgeschichte. Nichtsdestotrotz spiegeln wir vor das Verhalten des Getränkeautomaten mit einer sozio-historischen Eigenschaft (\$0,75 wert zu sein) des Objekts (der Münze), das das Verhalten verursacht, erklären zu können und vernachlässigen damit die intrinsi-

schen Eigenschaften, die in Wirklichkeit kausal relevant sind.³⁹² Diesen Umstand erklärt Dretske folgendermaßen:

We ignore them because, often enough, we don't even know what they are. Nonetheless, given the facts of supervenience, we know that, normally, inserting \$0,75 will get you a Coke even if we don't know which properties of the \$0,75 are responsible for this effect (is density relevant?). V is, after all, multiple realizable in S. We can use a variety of different coins, of different shapes and sizes, to make \$0,75. The machine will give us a Coke, it will behave in the same way, if we insert quarters, dimes and a nickel; or seven dimes and a nickel; or fifteen nickels. As long as the coins add up to \$0,75 we get the same result. So it is simpler and much more convenient in our explanations of machine behavior to mention the extrinsic V all the different Ss have in common even though we know it is S, not V, that explains the result. Convenience explains explanatory pretense.³⁹³

Die kausal relevanten, intrinsischen Eigenschaften werden also für gewöhnlich deshalb in der Erklärung ausgeklammert, weil sie häufig nicht bekannt sind. Aufgrund der Supervenienz des Geldwertes auf den intrinsischen Eigenschaften der Münzen ist es hingegen naheliegend, dass, wenn \$0,75 in den Automaten geworfen werden, normalerweise eine Cola ausgegeben wird, selbst falls unklar ist welche spezifischen Eigenschaften der \$0,75 genau für die Wirkung verantwortlich sind. Außerdem kann V von S multipel realisiert werden. D.h. es gibt einige Kombinationsmöglichkeiten von Münzen unterschiedlicher Größe und unterschiedlichen Gewichts, die zusammen \$0,75 ergeben. Ungeachtet der genauen Zusammensetzung des Kleingelds bleibt der Effekt (die Ausgabe einer Cola) aber stets derselbe, wenn die Geldstücke in ihrer Summe \$0,75 ergeben. Somit ist es wesentlich einfacher und bequemer, wenn in den Erklärungen des Verhaltens des Verkaufsautomaten auf den extrinsischen Wert V, den alle S gemeinsam haben, rekurriert wird, obwohl wir wissen, dass eigentlich S und nicht V das Ergebnis erklärt. Zusammenfassend folgert Dretske, dass die explanatorische Praxis also aus Gründen der Zweckdienlichkeit so ist, wie sie ist.

Nachdem Dretske sein Automaten-Beispiel detailliert beschrieben und den extrinsischen Charakter des Geldwerts in der Interaktion zwischen Metallstücken und Verkaufsautomaten herausgearbeitet hat, analysiert er wie diese Ergebnisse gemeinhin auf menschliches Verhalten übertragen werden:

There is a prevalent view in the philosophy of mind that the propositional attitudes (including belief) are something like internal coins. What you believe (intend, desire, conclude, regret, etc.) is an extrinsic property of the internal belief (intention, etc.) in the same way that the value of coins is extrinsic to the coins in the machine. For a materialist (who is not an eliminativist) a belief (some brain state, say) has intrinsic (neurobiological) properties, but it also has a content or meaning (= what it is one believes), and this is determined, in part at least, by the relations this internal state bears to external affairs. The relational individuation of belief is why the same belief can occur in much different heads; what makes it *that* belief is not the brain state that realizes it (this can be quite

³⁹² vgl. Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 262.

³⁹³ Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 262.

different), but the way that brain state is related to the rest of the world. Putnam's (1975) and Burge's (1979) examples have convinced many that, in this respect, beliefs are "internal money."
I will call this view the *Standard Theory*.³⁹⁴

In der Philosophie des Geistes ist die Auffassung sehr populär, dass propositionale Einstellungen, wie z.B. Glaubenssätze, so etwas wie ‚innere Münzen‘ sind. Dieser Vorstellung zufolge ist das, was man glaubt, wünscht, usw. eine extrinsische Eigenschaft von inneren Überzeugungen – also in etwa vergleichbar mit der Eigenschaft einer Münze einen bestimmten Geldwert zu besitzen. Für Materialisten, die nicht einer eliminativen Spielart dieser Theorie des Geistes anhängen, ist eine Überzeugung identisch mit einem Hirnzustand und besitzt somit sowohl intrinsische Eigenschaften (wie z.B. neurobiologische Attribute), als auch extrinsische Eigenschaften (etwa den Inhalt oder die Bedeutung des Geglaubten). Da aber die Bedeutung und der Inhalt von Überzeugungen zumindest teilweise durch Beziehungen zwischen dem inneren Zustand und der Außenwelt bestimmt werden, scheint es, als verhielten sich beispielsweise Überzeugung tatsächlich wie innere Münzen. Genauso wie verschiedene Geldstücke denselben Wert besitzen können, können auch verschiedene Personen denselben Gedanken haben. Daher spielt es für einen Gedanken auch eine größere Rolle in welchem Verhältnis er zum Rest der Welt steht als durch welchen Hirnzustand er genau realisiert ist – ein und derselbe Gedanke könnte in zwei Menschen auf sehr unterschiedliche Art und Weise verwirklicht sein. Diese Auffassung bezeichnet Dretske als Standardtheorie. Zudem ist Dretske überzeugt, dass kein Materialist beliebiger Couleur die Wahrheit der Standardtheorie anzweifeln könnte. Schließlich müssen Überzeugungen, die Verhalten kausal erklären, ein Teil des Systems sein, für das sie Erklärungen liefern. In anderen Worten müssen Überzeugungen auf irgendeine Weise physikalisch realisiert sein – z.B. über einen Hirnzustand, der bestimmte neurobiologische Qualitäten besitzt – denn sie erklären physikalisch beschreibbare Handlungen kausal. Andererseits ist dasjenige, was diesen inneren Zuständen ihren Inhalt und ihre Bedeutung verleiht, nicht im Kopf lokalisierbar. Dies ist vergleichbar mit dem Geldwert einer Münze, der gleichfalls nicht von einem physikalisch beschreibbaren Merkmal des konkreten Metallstücks, das die Münze ist, festgelegt wird. Generell sind Überzeugungen ihrem Wesen nach also repräsentational: was eine Überzeugung zu einer Überzeugung über Fußball anstatt über Philosophie macht, hängt davon ab in welcher Relation der innere Zustand zu äußeren Umständen steht. Denn es ist sicherlich nichts im Gehirn zu finden, dass eine beliebige neurobiologische Ereignis zu einer Überzeugung über Fußball macht und ein anderes zu einem Glaubenssatz über Philosophie. Der Inhalt respektive der repräsentationale Gehalt einer

³⁹⁴ Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 265f.

Überzeugung muss folglich in einem extrinsischen Verhältnis zum Eigner der Überzeugung stehen.³⁹⁵

Die vermeintlichen Konsequenzen, die ein Großteil von Philosophen aus der Standardtheorie ableitet, beschreibt Dretske wie folgt:

The Standard Theory is commonly thought to have the kind of epiphenomenal implications we uncovered in examining monetary-machine interactions. Although the content of a belief – *what* one believes – is routinely mentioned in explanations of behavior (just as the value of coins is mentioned in explanations of machine behavior), this content is, according to Standard Theory, as irrelevant to what we do as is the value of coins to what a machine does. If you want to know what makes vending machines dispense Cokes and candy bars, look to the intrinsic properties of internal causes – the shape, size, and weight of the internal coins that trigger its responses. For the same reason, if you want to know what makes people do the things they do, look not to the relational properties of belief (those that constitute *what* we believe) but to the intrinsic (i.e. neurobiological) properties of the belief. Look to the “shape” and “size” – that is, the syntax – of these internal “coins,” not their semantics. (...) Thus, the Standard View, while denying neither the reality nor the causal efficacy of the mental, leaves little or no room for understanding the causal efficacy of the mental qua mental. Beliefs, qua beliefs, have as much effect on the behavior of persons as do quarters, qua quarters, on the behavior of vending machines.³⁹⁶

Die Schlussfolgerung, die für gewöhnlich aus der Analogie zwischen dem Geldwert und dem Verhalten von Verkaufsautomaten auf der einen Seite und beispielsweise Überzeugungen und dem Verhalten von Personen auf der anderen Seite gezogen wird, besagt, dass Mentales in Erklärungen von Verhalten rein epiphänomenal zu verstehen ist. Denn, obwohl der Inhalt von Überzeugungen weithin in Erklärungen von Verhalten angeführt wird, ist er gemäß der Standardtheorie dennoch bedeutungslos für ein Verständnis der kausalen Grundlagen von Verhalten. Schließlich sind die eigentlichen Ursachen von Verhalten nicht in den relationalen Eigenschaften (wie z.B. Inhalt oder Bedeutung) von Hirnzuständen zu suchen, sondern vielmehr in deren neurobiologischen Charakteristika. Kurzum, die Standardtheorie leugnet weder die Realität noch die kausale Wirksamkeit von Mentalem, allerdings scheint es so als könne das Mentale unter diesen Umständen nicht qua Mentalem kausal wirksam sein, genauso wenig wie eine Münze qua ihres Geldwertes die Ausgabe eines Artikels aus einem Verkaufsautomaten verursachen kann. In beiden Fällen sind es jeweils *ausschließlich* die intrinsischen Eigenschaften der Ursachen kausal wirksam.

Im Zuge seiner Lösung dieses Problems weist Dretske an erster Stelle auf einen wichtigen Punkt hin, der in dessen Erörterung seines Erachtens bislang noch keine Beachtung gefunden hat:

We have, however, neglected an important aspect of the causal relations at work in monetary-machine and mind-body cases. In the monetary-machine interaction, for instance, there is the fact that the machines on which coins

³⁹⁵ vgl. Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 266.

³⁹⁶ Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 266f.

have a causal impact were designed and manufactured to be sensitive to objects having those intrinsic properties (S) on which monetary value supervenes, and they were made that way precisely *because* V supervenes on S. Business being what it was, machines that dispense commodities like cigarettes, food, and drink would not be designed to yield their contents to objects having S *unless* objects having S had V. Remove the fact of supervenience (as a result of widespread counterfeiting, say) and S objects will soon lose their causal power. They will no longer produce the effects they now produce. They will lose their causal power because machines will no longer be built to respond to objects having S. The causal efficacy of intrinsic S (on machines – not to mention people) depends on the supervenience of extrinsic V on S. Let V supervene on a different set of properties, T, and T-objects will, quickly enough, assume the causal powers of S-objects.³⁹⁷

Die Tatsache, die laut Dretske in der Diskussion um die Folgen der Analogie bis zu diesem Zeitpunkt völlig außer Acht wurde, bezieht sich auf den Kontext, in den die besagten Kausalbeziehungen eingebettet sind. So reagieren Verkaufsautomaten beispielsweise genau deshalb auf bestimmte intrinsische Eigenschaften S von Münzen, weil die Apparate in ihrer Planung und Herstellung so konzipiert wurden, dass sie auf eben diese Eigenschaften S ansprechen. Dies ist wiederum der Fall, weil das Haben eines bestimmten Geldwertes V auf dem Besitz der Eigenschaften S superveniert. Denn hätten die Objekte mit den Eigenschaften S nicht den Geldwert V, dann würden auch Verkaufsautomaten ihre Ware nicht ausgeben, sobald man Metallstücke mit den Eigenschaften S in sie einwerfen würde. Item, gäbe es die Supervenienz von V auf S nicht – das wäre beispielsweise dann der Fall, wenn sich hinreichend viele gefälschte Münzen mit den Eigenschaften S in Umlauf befänden – dann würden auch die Dinge, die die Eigenschaften S aufweisen, zeitnah ihre kausale Kraft verlieren. Schließlich würden die Hersteller von Verkaufsautomaten unter diesen Umständen ihre Geräte künftig sicherlich so konstruieren, dass sie nicht mehr auf Metallstücke mit den Eigenschaften S reagieren. Allgemein gesprochen bedeutet das, dass die kausale Wirksamkeit der intrinsischen Eigenschaften S von der Supervenienz der extrinsischen Eigenschaften V auf S abhängt. Angenommen, V würde fortan auf einer anderen Menge von Eigenschaften, z.B. den Eigenschaften T, supervenieren, dann würden T-Objekte gewiss in kürzester Zeit die kausale Wirksamkeit von S-Objekten übernehmen.

Dieser zusätzliche Aspekt der Kausalrelationen ist Dretske zufolge jedoch nicht gleichbedeutend damit, dass der Geldwert der eingeworfenen Münzen die Ausgabe von Artikeln aus dem Verkaufsautomaten auch *erklärt*. Vielmehr gibt der Automat seine Waren *de facto* deshalb aus, weil die eingeworfenen Metallstücke beispielsweise eine spezifische Größe, ein bestimmtes Gewicht und eine gewisse Dichte aufweisen. Allerdings sei nicht unbedingt die Frage warum ein Verkaufsautomat seine Ware ausgibt von Interesse, sondern vielmehr die Frage warum Objekte, die bestimmte Größen, Gewichte und Dichten haben, die Warenausgabe verursachen. In anderen Worten gilt es zu klären weshalb Objekte dieser Art Wirkungen jener

³⁹⁷ Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 270f.

Art verursachen. Dies wiederum hängt nach Dretske zumindest zum Teil von einer starken Korrelation zwischen Dingen mit charakteristischen Größen, Gewichten und Dichten einerseits und deren Besitz eines bestimmten Geldwertes andererseits zusammen. Folglich spielt die schwache Supervenienz von V auf S eine Schlüsselrolle in diesem Zusammenhang. Der Geldwert von Münzen erklärt also nicht weshalb Artikel von Automaten ausgegeben werden, aber er erklärt weswegen Geldstücke den Auswurf von Ware verursachen.³⁹⁸

Überträgt man diese Argumentationsstrategie auf Fälle mentaler Verursachung kommt man zu folgendem Ergebnis:

If we think of ourselves as “vending machines” whose internal causal structure is designed, shaped, and modified not, as with vending machines, by engineers, but, in the first instance, by evolution and, in the second, by learning, then we can say that although it is the “size” and “shape” (the syntax, as it were) of the internal causes that make the body move the way it does (just as it is the size and shape of the coins that releases the Cokes) it is, or may be, the fact that a certain extrinsic property supervenes on that neurological “size” and “shape” that explains why internal events having these intrinsic properties have the effect on the body they have. What explains why a certain neurological event in the visual cortex of a chicken – an event caused by the shadow of an overhead hawk – causes the chicken to cower and hide is the fact that such neurological events have a significant (to chickens) extrinsic property – the property of normally being caused by predatory hawks. It is, or may be, possession of this extrinsic property – what the internal events *indicate* about external affairs – that explains why objects having those intrinsic properties cause what they do.³⁹⁹

Genauso wie die innere kausale Struktur von Automaten durch die für ihre Entwicklung verantwortlichen Ingenieure festgelegt wurde, wird die innere kausale Struktur des Menschen in erster Linie durch die Evolution und in zweiter Linie durch Lernprozesse determiniert. Die Ursache dafür, dass sich der Körper in der Art und Weise bewegt wie er sich bewegt, ist sozusagen in der Syntax (i.e. in der konkreten neurologischen ‚Verdrahtung‘) der inneren Ursachen zu finden. Allerdings könnte die Tatsache, dass eine gewisse extrinsische Eigenschaft auf den ‚harten‘ neurologischen Parametern der inneren Ursache superveniert, erklären warum gerade diejenigen innere Ursachen, die exakt diese intrinsischen Eigenschaften besitzen, genau jene Wirkung auf den Körper ausüben, die sie tatsächlich haben. Zur Veranschaulichung gibt Dretske folgendes Beispiel: Ein Huhn sieht den Schatten eines Falken, der gerade über es hinweg fliegt, woraufhin es sich duckt und versteckt. Der Grund weswegen ein neurologisches Ereignis im visuellen Kortex des Huhns erklärt warum sich das Huhn duckt und versteckt liegt in der Tatsache, dass solche neurologischen Ereignisse eine für Hühner lebensbedrohliche extrinsische Eigenschaft besitzen – nämlich die Eigenschaft von Falken, also von Fressfeinden der Hühner, verursacht worden zu sein. Daher könnte also der Besitz einer extri-

³⁹⁸ vgl. Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 271.

³⁹⁹ Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 271.

nischen Eigenschaft durchaus erklären, weshalb Objekte mit gewissen intrinsischen Eigenschaften das verursachen, was sie verursachen.

Von dieser Aussage zur Schlussfolgerung, dass eigentlich die extrinsischen Eigenschaften der inneren Ursachen und nicht deren intrinsische Pendanten Verhalten kausal erklären, ist es nur ein kleiner Argumentationsschritt. Das einzige Element, das benötigt wird um diesen Gedankengang zu vollenden ist lediglich die Prämisse, dass unter ‚Verhalten‘ nicht die Körperbewegungen zu verstehen sind, die von den inneren Ereignissen verursacht werden, sondern vielmehr die Verursachung dieser Bewegungen durch innere Ereignisse. In anderen Worten bedarf es einer klaren begrifflichen Unterscheidung zwischen ‚*Verhalten*‘, dieses wird von bestimmten Überzeugungen erklärt, und ‚*Körperbewegungen*‘, aus denen das Verhalten zum Teil besteht. Wenn also mein Bewegen meiner Arme und Beine nicht mit den Bewegungen meiner Extremitäten gleichzusetzen ist, sondern vielmehr den Umstand beschreibt, dass ein Ereignis in mir die Regungen meiner Körperteile verursacht, dann gilt Folgendes: die intrinsischen, d.h. neurologischen, Eigenschaften des Ereignisses in meinem Körper erklären die konkreten Körperbewegungen, die ich vollziehe. Die extrinsischen Charakteristika des Ereignisses in meinem Inneren stehen wiederum mit den äußeren Bedingungen, mit denen die intrinsischen Eigenschaften korrelieren, in Kontakt und erklären somit warum ich mich bewege. Offensichtlich hängt Dretskes Ansatz zur Lösung des Externalismus-Problems davon ab, dass bei der Erklärung von Verhalten durch Gründe eine Korrelation zwischen den extrinsischen und intrinsischen Eigenschaften der Gründe besteht. Genauer gesagt, muss eine schwache Supervenienz der extrinsischen auf den intrinsischen Eigenschaften gegeben sein.⁴⁰⁰

Zu den Fragen, ob eine Analogie zwischen dem Verhalten von Münzautomaten und dem Verhalten von Menschen besteht und ob ein solches Verhältnis das Mentale zu einem Epiphänomen macht, zieht Dretske ein abschließendes Fazit:

Yes, beliefs stand to human behavior in something like the way money stands to vending machine behavior. Does this show that what we believe is causally irrelevant to what we do? No, it does not show this any more than it shows that the fact that nickels, dimes, and quarters have monetary value is irrelevant to the behavior of vending machines. The fact that these coins have monetary value, that fact that they are a widely accepted medium of exchange, explains why the machines (are built to) dispense their contents when objects of this sort are placed in them. In this sense, the fact that these coins have monetary value explains why machines behave the way they do when coins are in them. The same is true of belief: the extrinsic properties of these beliefs – *what it is we believe* – explains why we behave the way we do when these beliefs occur in us.⁴⁰¹

⁴⁰⁰ vgl. Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 272.

⁴⁰¹ Dretske, Fred: „*Minds, Machines, and Money: What Really Explains Behavior*“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000, S. 272f.

Die Analogie zwischen dem Verhalten von Verkaufsautomaten und menschlichem Verhalten hat Dretske zufolge durchaus Bestand, denn Überzeugungen stehen im selben Verhältnis zum Verhalten von Personen wie Geld zum Verhalten von Münzautomaten. Das hat jedoch nicht zur Folge, dass das, was jemand glaubt, kausal irrelevant für sein Handeln wäre. Genauso wenig kann man davon ausgehen, dass das Haben eines Geldwerts von Münzen kausal irrelevant für das Verhalten von Automaten sei. Schließlich erklärt die Tatsache, dass Münzen einen bestimmten Geldwert besitzen, warum Verkaufsautomaten so konstruiert sind, dass sie ihre Ware ausgeben, sobald Metallstücke mit den physikalischen Eigenschaften von Münzen in sie eingeworfen werden. Aus diesem Grund erklärt die Tatsache, dass Münzen einen spezifischen Geldwert besitzen, weshalb Münzautomaten auf den Einwurf von Geld in der Weise reagieren wie sie es für gewöhnlich tun. Die Parallele zu Überzeugungen ist, dass auch hier die extrinsischen Eigenschaften – in anderen Worten der Inhalt der Überzeugungen – erklären warum man sich auf diese Weise verhält, wenn man jene Überzeugungen besitzt.

Schluss

Abschließend stellt sich natürlich die Frage, welcher Theorie es gelingt die momentan beste Antwort auf die bisher beschriebenen Probleme bei der Erklärung mentaler Verursachung zu geben. Meiner Meinung nach ist es Fred Dretskes Ausformulierung einer Dual-Explanandum Theorie der geeignetste Kandidat dafür, die diskutierten Schwierigkeiten einer kohärenten Darstellung mentaler Verursachung auszuräumen. In einem Satz zusammenfasst behauptet Dretske, dass mentale und physikalische Eigenschaften in Kausalerklärungen nicht auf denselben Erklärungsgegenstand referieren, anders ausgedrückt weisen sie zwei unterschiedliche Explanada auf. So erklären mentale Eigenschaften beispielsweise das Verhalten von Personen, während physikalische Eigenschaften wiederum Erklärungen für Körperbewegungen liefern, so Dretske.

Auslösende und strukturierende Ursachen

Ein wesentlicher Schritt in seinem Argumentationsgang ist die Unterscheidung von zwei Arten von Ursache: *auslösenden Ursachen* (*triggering causes*) auf der einen Seite und *strukturierenden Ursachen* (*structuring causes*) auf der anderen. Diese Begriffsdifferenzierung führt Dretske in seinem bahnbrechenden Aufsatz *Mental Events as Structuring Causes of Behaviour* ein. An dessen Anfang stellt er die Aussage, kausale Erklärungen seien grundsätzlich kontextabhängig. So spiele es beispielsweise eine essentielle Rolle für die Auswahl einer geeigneten Ursache⁴⁰² von E, welche Interessen mit der Kausalerklärung von E verfolgt werden, welchen Zwecken gedient werden soll und welches Wissen bereits vorhanden ist. Schließlich hänge nahezu jedes beliebige Ereignis E von einer Vielzahl von anderen Ereignissen in einer Art und Weise ab, die jedes einzelne dieser Vorkommnisse – je nach Kontext – als Ursache von E im Rahmen einer Kausalerklärung qualifiziert. Die Unterscheidung von strukturierenden und auslösenden Ursachen in Kausalerklärungen des Verhaltens eines Systems, ist laut Dretske im Wesentlichen eine Differenzierung zwischen psychologischen und biologischen Erklärungen des Verhaltens.⁴⁰³

⁴⁰² Diese Redeweise ist etwas irreführend, denn genau genommen vertritt Dretske nicht die Ansicht, dass ein Ereignis eine bestimmte Anzahl miteinander konkurrierender Ursachen besitzen kann. Ihm geht es vielmehr um Kausalerklärungen für ein Ereignis und diesbezüglich konstatiert er eine gewisse Ambivalenz beim Heranziehen einer für die entsprechende Erklärung geeigneten Ursache.

⁴⁰³ Vgl. Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 121.

Den Unterschied zwischen strukturierenden und auslösenden Ursachen illustriert Dretske anhand folgenden Beispiels:

An operator moves the cursor on a screen by pressing a key on the keyboard. Pressure on the key is the triggering cause of cursor movement. The movement would not have occurred if the key had not been pressed. Pressure on this key makes the cursor move. In conditions that exist at the time the key is pressed, pressure on the key (this type of event) is regularly followed by cursor movement (that type of event). It is this kind of causal arrangement, the sort characteristic of a triggering cause, that allows us to speak of the operator himself as moving the cursor by pressing the key.

On the other hand, we sometimes speak of the events that produced hardware conditions (actual electrical connections in the computer) and programming (software) as the cause of cursor movement. This is especially evident when cursor movement (in response to pressure on a certain key) is unexpected or unusual. Imagine a puzzled operator, watching the cursor move as he pokes the key, asking: 'Why is the cursor moving?' Since the operator knows that the pressure on the key is making the cursor move (that, in fact, is what he finds puzzling), a different explanation of cursor movement is being sought. The operator is looking for the structuring cause. He wants to know what brought about or caused the machine to occupy a state, to be in a condition, in which pressure on the key has this effect. He knows, or can easily be assumed to know (after a few presses of the key) that *E* (cursor movement) is being caused by *T* (pressure on the key). What he wants to know is why it is. Who or what made *E* depend on *T* in this way?⁴⁰⁴

In Dretskes Beispiel drückt ein Benutzer eine Taste auf einer Computertastatur, woraufhin sich der Cursor auf dem Bildschirm bewegt. Der Tastendruck ist die auslösende Ursache für die Cursorbewegung. Wäre die entsprechende Taste nicht betätigt worden, dann hätte sich auch die Eingabemarkierung auf dem Monitor nicht bewegt. Grundsätzlich folgen unter den, zum Zeitpunkt des Drückens auf die Taste gegebenen Umständen, die Ereignisse der einen Kategorie (Cursorbewegung) stets auf Ereignisse, die der anderen Klasse (Tastendruck) zugehörig sind. In der Alltagssprache wird dieser, für auslösende Ursachen typische Umstand, wiedergegeben, indem man davon spricht, dass der Computernutzer *selbst* den Cursor bewegt hat. Allerdings ist diese Kausalerklärung für die Cursorbewegung nicht die einzige, die regelmäßig Verwendung findet. Denn in anderen Kontexten bezieht man sich mit der Wendung ‚die Ursache der Cursorbewegung‘ entweder auf die mechanische und elektrische Implementierung der Cursorbewegung oder auf die Programmierung, die sich dafür verantwortlich zeichnet – also auf die konkreten Hardware- bzw. Softwarezustände, die die Cursorbewegung hervorbringen. Diese Form der kausalen Erklärung der Cursorbewegung erfolgt in der Regel insbesondere dann, wenn die Cursorbewegung für den Benutzer unerwartet eintritt. So ist folgendes Szenario beispielsweise durchaus lebenssecht: ein Computernutzer beobachtet überrascht wie sich der Cursor bewegt nachdem er eine bestimmte Taste gedrückt hatte. Erstaunt fragt er, warum sich der Cursor bewegt hat. Da ihm – zumindest nach einigen Versuchen – offensichtlich klar sein dürfte, dass der Druck auf die Tastatur *T* die Cursorbewegung *E* hervorgerufen hat, sucht er nach einer anders gearteten Kausalerklärung für die Cursorbewe-

⁴⁰⁴ Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 121f.

gung. Letzten Endes sucht er nach einer strukturierenden Ursache. Das heißt, er möchte wissen worin die Ursache dafür liegt, dass sich der Computer im Moment des Tastendrucks in einem Zustand befunden hat, in dem der Druck auf diese Taste jene Wirkung hervorgerufen hat. Anders aufgedrückt möchte der Nutzer also nicht wissen, ob T E verursacht hat (dessen ist er sich nämlich bewusst – dies hatte ursprünglich seine Verwunderung hervorgerufen), vielmehr ist er interessiert zu erfahren, warum dies der Fall ist. Worauf oder auf wen ist es zurückzuführen, dass E in dieser Art und Weise von T abhängt.

Ein weiteres Beispiel zur Verdeutlichung seiner Unterscheidung zwischen auslösenden und strukturierenden Ursachen liefert Dretske mit folgendem Szenario:

A terrorist plants a bomb in the general's car. The bomb sits there for days until the general gets in his car and turns the key to start the engine. The bomb is detonated (triggered by turning the key in the ignition) and the general is killed. Who killed him? The terrorist, of course. How? By planting a bomb in his car. Although the general's own action (turning on the engine) was the triggering cause, the terrorist's action is the structuring cause, and it will be his (the terrorist's) action, something he did a week ago, that will certainly be singled out, in both legal and moral inquiries, as the cause of the explosion that resulted in death.⁴⁰⁵

Dretskes zweites Beispiel handelt von einem Terroristen, dem es gelingt eine Autobombe im Wagen eines Generals zu installieren. Einige Tage nachdem der Attentäter den Sprengsatz im Fahrzeug des Offiziers deponiert hat, besteigt der Soldat den PKW und bringt die Bombe zur Detonation, indem er den Zündschlüssel umdreht. Obwohl nun eine Handlung des Generals – das Anlassen des Motors – seinen Tod (in Sinne einer auslösenden Ursache) verursacht hat, würde man dennoch behaupten, der Terrorist habe dessen Ableben (im Sinne einer strukturierenden Ursache) verursacht, indem er den Zündmechanismus des Wagens mit einem Sprengkörper verbunden hat. Untersucht man schließlich den Exitus des Militärs vor einem juristischen oder moralischen Hintergrund, so ist es auch die strukturierende Ursache, die maßgeblich als Ursache der Explosion und des Todes betrachtet wird.

Offensichtlich haben die Angabe einer auslösenden respektive einer strukturierenden Ursache auch unterschiedliche Arten von Kausalerklärungen zur Folge. Denn die Art von Beziehung, die diese beiden Gattungen von Ursachen zu ihren jeweiligen Wirkungen unterhalten, ist höchst unterschiedlich. Die Relation zwischen auslösenden Ursachen und ihren Wirkungen steht sozusagen in der Tradition eines Humeschen Kausalitätsverständnisses als konstante Konjunktion zweier Ereignistypen. Schließlich gilt auch für das Hervorbringen einer Wirkung E durch eine auslösende Ursache T folgende Formel: unter den allgemeinen Bedingungen, die gegeben waren, als sich T ereignet hat, folgen Ereignisse, die zum Typ E gehören in der Regel auf Ereignisse, die dem Typ T zuzurechnen sind. Bei strukturierenden Ursachen hingegen

⁴⁰⁵ Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 122f.

ist dies nicht der Fall. Das heißt, im Falle strukturierender Ursachen sind das gemeinsame Auftreten der Umstände, unter denen sich S ereignet hat, und S selbst nicht bereits allein zwangsläufig hinreichend für das Eintreten von E. Denn für strukturierende Ursachen ist es essentiell wichtig, dass spätere Ereignisse, die von S unabhängig sind, eintreten, um S in eine strukturierende Ursache von E umzuwandeln. Allerdings liegt es außerhalb der Einflussosphäre von S, ob diese Ereignisse tatsächlich stattfinden oder nicht. Somit existieren zwischen strukturierenden Ursachen (im Gegensatz zu auslösenden Ursachen) und deren Wirkungen keine Regelmäßigkeiten der Form: Immer wenn S unter den gegebenen Umständen eintritt, ereignet sich E. Hierzu ein Beispiel: Ein Computerexperte verkabelt einen Rechner neu, so dass der Druck auf eine bestimmte Taste eine Cursorbewegung zur Folge hat. Es entscheidet sich jedoch niemals jemand dazu diese Taste zu drücken. Nichtsdestotrotz wäre die Handlung des Computerexperten in einem Szenario, in dem ein Nutzer die entsprechende Taste drückt, die strukturierende Ursache der Cursorbewegung gewesen. Ob ein Ereignis tatsächlich zu einer strukturierenden Ursache wird, unterliegt also im Wesentlichen dem zukünftigen Verlauf der Welt.⁴⁰⁶

Ein weiteres wichtiges Unterscheidungsmerkmal zwischen auslösenden und strukturierenden Ursachen besteht im Beziehungstyp der Kausalrelationen: während strukturierende Kausalbeziehungen eine eins-zu-viele-Relation repräsentieren, gehören auslösende Kausalbeziehungen dem Typ der eins-zu-eins-Relation an. Auf das Cursorbeispiel bezogen heißt das, dass die auslösende Ursache für jede Cursorbewegung eine andere ist, wohingegen die strukturierende Ursache immer dieselbe bleiben könnte. Nehmen wir beispielsweise an, ein Computerspezialist programmiert einen Rechner so, dass der Druck auf eine bestimmte Taste eine Cursorbewegung hervorbringt. Nun drücken beispielsweise zwei Nutzer jeweils die entsprechende Taste zu unterschiedlichen Zeitpunkten. In diesem Fall wären die jeweiligen auslösenden Ursachen der Cursorbewegung verschieden (i.e. Druck auf Taste von Nutzer 1 bzw. Tastendruck durch Nutzer 2), die strukturierende Ursache ist jedoch für beide Begebenheiten identisch. Der Computerfachmann muss also nicht vor jedem einzelnen Tastendruck den Computer neu programmieren, eine einzige Umprogrammierung genügt.⁴⁰⁷

Der folgende Einwand gegen eine Einteilung in auslösende und strukturierende Ursachen liege offensichtlich auf der Hand, so Dretske:

⁴⁰⁶ Vgl. Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 123.

⁴⁰⁷ Vgl. Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 123f.

Some may object to this way of describing things. They may prefer to say that what I am calling the structuring cause of *E* is not a cause of *E* at all. It is, rather, a cause (a good, old-fashioned cause) of those background or standing conditions (call them *B*) in which *T* causes *E*. So instead of having two different types of cause (triggering and structuring) for *E*, we have one sort of cause, a triggering cause, for different effects: *T* (what I am calling the triggering cause of *E*) causes *E*, but *S* (what I am calling the structuring cause of *E*) causes *B*, the conditions (or one of the conditions) in which *T* causes *E*. So, if you like, is the cause, not of *E*, but of *T*'s causing *E*.⁴⁰⁸

Dretskes Klassifikation, derzufolge zwei unterschiedliche Arten von Ursachen existieren, könnte also durch folgende Behauptung zurückgewiesen werden: de facto gibt es nicht zwei verschiedene Typen von Ursachen, sondern im Grunde treten Ursachen nur in einer einzigen Erscheinungsform auf – und zwar in Form von auslösenden Ursachen. Was Dretske als strukturierende Ursache *S* von *E* bezeichnet, ist im eigentlichen Wortsinn gar keine Ursache von *E*. Vielmehr handelt es sich bei *S* um eine (auslösende) Ursache der Hintergrundbedingungen *B*, vor denen *T* *E* verursacht. Somit besitzt *E* in Wirklichkeit nicht zwei verschiedene Ursachen – eine auslösende und eine strukturierende, sondern tatsächlich existiert nur der Typus der auslösenden Ursachen. Kurzum, *T* ist die auslösende Ursache von *E*, während *S* die auslösende Ursache von *B* ist, oder, in anderen Worten, *S* ist die Ursache von *T*'s Verursachen von *E*.

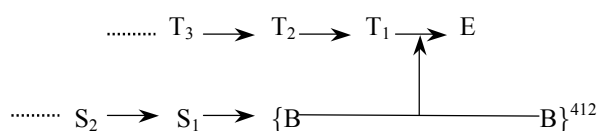
Grundsätzlich hat Dretske keine Einwände gegen diese Art der Darstellung. Im Gegenteil, es sei für bestimmte Fälle sogar durchaus angebracht eine strukturierende Ursache von *E* nicht als Ursache von *E*, sondern als Ursache der Hintergrundbedingungen *B*, unter denen ein anderes Ereignis (eben die auslösende Ursache *T*) *E* hervorbringt, zu betrachten. Von diesem Standpunkt aus gesehen, verursachen beispielsweise Computerprogrammierer nicht einzelne Cursorbewegungen – diese werden vielmehr von den jeweiligen Nutzern herbeigeführt. Die Softwareentwickler sorgen genauer gesagt dafür, dass sich der Rechner in einem Zustand befindet, in dem der Druck auf eine bestimmte Taste durch einen Computerbenutzer eine Cursorbewegung hervorruft. Demnach ist eine strukturierende Ursache von *E* eigentlich eher als eine Ursache der mehr oder weniger stabilen Hintergrundbedingungen *B* zu verstehen. Diese wiederum gewährleisten, dass Ereignisse des Typs *E* von Ereignissen des Typs *T* in einer Art und Weise abhängen, dass Ereignistoken von *T*, falls und sobald sie eintreten, Token von *E* verursachen. Obwohl Dretske mit dieser Darstellung des Zusammenhangs von *T*, *S* und *E* einverstanden ist, hält er es dennoch für gerechtfertigt weiterhin von *S* als einer strukturierenden Ursache von *E* zu sprechen.⁴⁰⁹ Schon allein „(...) out of deference to those (and I think this is most of us most of the time) who think that a cause of *E* is an earlier event on which *E* is counterfactually dependent in the right way, a way which allows us (given a suitable con-

⁴⁰⁸ Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 124.

⁴⁰⁹ Vgl. Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 124f.

text) to single it out as the cause of *E* in causal explanations of *E*.“⁴¹⁰ Die Redeweise von *S* als (strukturierender) Ursache von *E*, sei also bereits dadurch gerechtfertigt, dass nach allgemeinem Verständnis eine Ursache von *E* ein Ereignis ist, von dem *E* kontrafaktisch auf die richtige Art und Weise abhängt. Diese korrekte Form kontrafaktischer Abhängigkeit ist dadurch gekennzeichnet, dass es möglich sein muss dieses Ereignis in Kausalerklärungen von *E* als Ursache von *E* identifizieren zu können. So hätte sich zum Beispiel der Cursor (nachdem der Computernutzer eine bestimmte Taste gedrückt hatte) nicht bewegt, wenn die Hard- oder Software des Computers vorher nicht entsprechend eingerichtet worden wäre. Oder der General wäre immer noch am Leben, wenn der Terrorist nicht die Bombe in dessen Auto installiert hätte. Im alltäglichen Sprachgebrauch werden die Eingriffe in die Programmierung oder in die elektrische Verdrahtung eines Rechners durch Computerexperten, sowie der Einbau einer Bombe ins Fahrzeug des Generals kurzerhand als Ursachen der Cursorbewegung bzw. des Todes des Generals betrachtet. Dies reiche laut Dretske bereits aus diesen Ereignissen den Status einer Ursache zuzuschreiben. Folglich schlägt Dretske vor: „(...) I propose to continue speaking of events which ‘configure’ circumstances so as to make (tokens of) *T*, when (and if) they occur, cause (tokens of) *E*, events which (in this sense) cause *T* to cause *E*, as themselves causes of *E*.“⁴¹¹ Generell ist es nach Dretske also auch legitim Ereignisse der Kategorie *S*, die die Umstände derart ‚gestalten‘, dass Ereignistoken des Typs *T*, wenn sie eintreten, Token des Ereignistyps *E* verursachen, selbst auch als Ursachen von *E* zu bezeichnen. Allerdings beinhalte diese Sprechweise auch die Gefahr einer Verwechslung von Ursachen von verschiedenen Arten von Dingen mit unterschiedlichen Arten von Ursachen desselben Dings – solange man sich jedoch dessen bewusst ist, was mit strukturierender Ursache gemeint ist und worin sich eine solche von einer auslösenden Ursache unterscheidet, sei dies unproblematisch, so Dretske.

Die graphische Darstellung seiner Konzeption von strukturierenden und auslösenden Ursachen, die Dretske in seinem Aufsatz vorstellt, sieht folgendermaßen aus:



⁴¹⁰ Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 125.

⁴¹¹ Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 125.

⁴¹² Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 126.

Dazu erläutert er:

For any event E that has a causal explanation, there exists a chain of triggering causes stretching from the proximal, T_1 , to the remote ($T_2 \dots$). There is also a chain of structuring causes – once again going from proximal, S_1 , to the remote – for the condition, B , on which the causal relation between T_1 and E depends. The vertical arrow between B and the causal relationship (\rightarrow) between T_1 and E is meant to signify this dependency relation: S_1 , we might say (and in *Explaining Behavior* I often did say), causes T_1 to cause E by creating the conditions (or one of the conditions) on which the causal relationship between T_1 and E depends. B is a condition whose onset necessarily occurs before T_1 and whose termination is (typically) after E .⁴¹³

Laut Dretske existiert also für jedes Ereignis E eine Kette von auslösenden Ursachen, die sich von nahen auslösenden Ursachen wie T_1 hin zu weiter entfernten auslösenden Ursachen (etwa T_2 , usw.) erstreckt.⁴¹⁴ Ganz analog verhält es sich mit strukturierenden Ursachen – auch hier besteht eine Aufeinanderfolge von strukturierenden Ursachen (von kurz zurückliegenden wie S_1 zu bereits weiter in der Vergangenheit zu verortenden, wie zum Beispiel S_2 , etc.), die sich für die Hintergrundbedingungen B , von denen die Kausalbeziehung zwischen T_1 und E abhängt, verantwortlich zeichnen. Der senkrechte Pfeil zwischen B und der Kausalbeziehung ($T_1 \rightarrow E$) wiederum, kennzeichnet Dretske zufolge eine Abhängigkeitsrelation, die folgende Redeweise legitimiert: S_1 verursacht, dass T_1 E verursacht, indem es die Bedingungen (oder zumindest eine Bedingung) erschafft, von denen die Kausalrelation zwischen T_1 und E abhängig ist. Die besagten Hintergrundbedingungen B sind nicht mit einem punktuellen Ereignis, das zu einem genau bestimmbar Zeitpunkt eintritt, vergleichbar, sondern erstrecken sich vielmehr in der Regel über eine gewisse Zeitspanne. Dieser Zeitraum beginnt notwendigerweise vor dem Eintreten von T_1 und endet üblicherweise nachdem sich E ereignet hat. Offensichtlich besteht in diesem Punkt ein wesentlicher Unterschied zwischen auslösenden und strukturierenden Ursachen. Während nämlich eine Kausalkette mit ausschließlich auslösenden Ursachen eine klare chronologische Abfolge besitzt, d.h. für jedes Glied in der Kette gilt, dass es erst dann und nur dann eintritt, wenn das unmittelbar vorausgehende bereits beendet ist, ist dies bei strukturierenden Ursachen augenscheinlich nicht notwendigerweise der Fall. Aufgrund dieser Besonderheiten fallen strukturierende Ursachen bei der Suche nach Ursachen oft durch das Raster. Hierzu Dretske:

⁴¹³ Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 126.

⁴¹⁴ Aus Gründen der Komplexitätsreduzierung beschränkt sich Dretske in seinem Beispiel auf je eine einzige Kausalkette auslösender und strukturierender Ursachen. Tatsächlich ist es aber durchaus vorstellbar, dass mehrere Ketten von auslösenden Ursachen in ein und demselben Ereignis konvergieren, wobei jede dieser Ketten, sowie jedes einzelne Element in der Kette, wiederum eine eigene strukturierende Ursache aufweisen kann. Zudem könnte es auch beispielsweise der Fall sein, dass verschiedene strukturierende Ursachen für die einzelnen Bedingungen, von denen die Kausalbeziehung zwischen auslösender Ursache und Wirkung abhängt, existieren. So könnte die Hintergrundbedingung B de facto aus einer Menge bestimmter Unterbedingungen B_1 , B_2 , etc. zusammengesetzt sein, von denen wiederum jede eine unterschiedliche Ursache besitzt.

B itself is a persisting condition or state of affairs, not really the sort of event-like occurrence or happening that we are accustomed to picking out as a cause. Besides (...), *B* typically extends beyond the effect (*E*) for which we are seeking a cause. As a result, in picking out structuring causes, *B* is generally disqualified (or simply ignored) in favour of those events (typically actions) that occur (wholly) before *E* and help bring it about by creating the conditions (*B*) on which the triggering of *E* depends. Hence, the most recent structuring cause can occur in the very remote past. *S_j*, the most proximal (recent) structuring cause of *E*, may have occurred years before its effect. What transmits its influence, so to speak, is the persisting background condition it brought about (*B*).⁴¹⁵

Die Gründe weshalb der, von strukturierenden Ursachen hervorgerufenen, Hintergrundbedingung *B* selbst der Status einer ‚echten‘ Ursache oftmals verwehrt wird, sieht Dretske im Wesentlichen in zwei ihrer Eigenschaften: erstens findet sie nämlich nicht wie ein gewöhnliches Ereignis zu einem bestimmten, klar eingegrenzten Zeitpunkt statt, sondern besteht als fortwährender Zustand, gegebenenfalls sogar über einen sehr langen Zeitraum hinweg. Zweitens ist es zudem typisch für die Hintergrundbedingung *B*, dass sie auch nach dem Eintreten des Wirkungsereignisses *E*, fort dauert (dies ist jedoch nicht zwangsläufig der Fall, vgl. Dretskes Terroristen-Beispiel). Diese beiden Aspekte der Hintergrundbedingung *B* führen dazu, dass bei der Auswahl einer strukturierenden Ursache, Ereignisse – insbesondere Handlungen – bevorzugt werden, die vor dem Eintreten von *E* stattgefunden haben und die an der Schaffung des Umstands *B*, von denen die auslösende Ursache von *E* abhängt, beteiligt waren. Daher kann es zum Beispiel auch der Fall sein, dass selbst die zeitlich am nächsten an *E* liegende strukturierende Ursache *S₁*, weit in der Vergangenheit liegt. Auf das Terroristen-Beispiel bezogen, könnte etwa der Attentäter die Bombe bereits vor Monaten oder gar vor Jahren im Wagen des Generals platziert haben (*S₁*). Trotzdem wäre der kausale Einfluss der Handlung des Saboteurs dadurch gesichert, dass hiermit die Bedingung *B* erschaffen wurde, die gemeinsam mit dem Betätigen der Zündung durch den hochrangigen Militär (*T₁*) zu der Explosion seines Autos und infolgedessen zu seinem Tod (*E*) führt. Der kausale Einfluss der strukturierenden Ursache *S₁* wird sozusagen durch deren Hervorbringen der beständigen Hintergrundbedingungen *B* auf die Wirkung *E* – also die Sprengung des Wagens und das Ableben des Generals – übertragen.

Um seine Beschreibung der Unterschiede zwischen auslösenden und strukturierenden Ursachen abzuschließen, weist Dretske schließlich noch auf Fälle hin, bei denen die beiden Ursachentypen zusammenzufallen scheinen. Ein solcher Fall ist beispielsweise dieser:

Suppose a dim-witted terrorist forgets he planted a bomb. Or he forgets which car it was in which he planted the bomb. A few days later, needing a car, he steals the ‘wired’ car and blows himself up. Is the terrorist both triggering and structuring cause of his own death? He created the conditions that enabled him to (unintentionally) blow himself up, yes, but that only means that one and the same person was involved in both causes. It doesn’t show that the causes are the same. What he did to trigger this outcome is different from what he did to structure

⁴¹⁵ Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 127.

it. It was his turning on the ignition that triggered the explosion; it is was his wiring the bomb to the ignition (a week ago) that structured it. Though both events (actions) involve the terrorist, they are quite different.⁴¹⁶

Eine kleine Modifikation des Terroristen-Beispiels erschafft ein Szenario, in dem auslösende und strukturierende Ursache auf den ersten Blick zu verschwimmen scheinen. Angenommen, ein minderbemittelter Terrorist platziert eine Bombe in einem Auto und vergisst dies anschließend bzw. kann sich nicht mehr daran erinnern, in welchem Wagen er genau den Sprengkörper installiert hat. Einige Tage später benötigt der Attentäter ein Fahrzeug und stiehlt zufälligerweise den Pkw, den er zuvor mit einer Bombe versehen hat. Als er den Anlasser betätigt, explodiert das Auto. Nun könnte man der Auffassung sein, der Terrorist sei in diesem Beispiel sowohl die auslösende als auch die strukturierende Ursache seines eigenen Todes, bzw. dass, in diesem Szenario beide Ursachentypen zusammenfallen. Dies ist laut Dretske jedoch nicht der Fall, selbst wenn man intuitiv zunächst geneigt sein sollte ein solches Urteil zu fällen. Zwar hat der Attentäter sowohl an der auslösenden als auch an der strukturierenden Ursache seines eigenen Todes einen wesentlichen Anteil, dennoch sind beide Ursachen nicht deckungsgleich. Schließlich sind die Handlungen des Terroristen, die seinen Tod auslösen, nicht die selben Aktionen, die sein Ableben strukturieren. Während nämlich sein Umdrehen des Zündschlüssels die Explosion und somit den Exitus des Terroristen ausgelöst hat, ist es sein Verkabeln des Autos mit dem Sprengsatz, das es strukturiert hat. Obwohl er also an beiden Ereignissen bzw. Handlungen konstitutiv beteiligt war, sind diese offensichtlich nicht identisch. Per Analogieschluss überträgt Dretske diesen Gedankengang auf Eigenschaften von Dingen: „Just as different actions of the same person can operate as triggering and structuring causes, different states of the same object (a single object having different properties) can function the same way. *O*'s possession of *P*₁ might be the triggering cause of *E* while its possession of some quite different property, *P*₂, say, can be the structuring cause of the same effect.“⁴¹⁷ Genauso wie unterschiedliche Handlungen ein und derselben Person die auslösende und die strukturierende Ursache eines Ereignisses verkörpern können, ist dies für unterschiedliche Zustände ein und desselben Objekts möglich. Letzteres ist wiederum lediglich eine andere Redeweise dafür, um auf unterschiedliche Eigenschaften einer bestimmten Entität Bezug zu nehmen. So kann beispielsweise das Haben der Eigenschaft *P*₁ eines Objekts *O* die auslösende Ursache für *E* sein, wohingegen in seinem Besitz einer anderen Eigenschaft, etwa *P*₂, die strukturierende Ursache von *E* zu verorten sein könnte. Überträgt man diese Vorstellung auf Ereignisse im Gehirn, mittels derer menschliches Verhalten erklärt werden soll,

⁴¹⁶ Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 128.

⁴¹⁷ Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 128.

dann könnte beispielsweise der Besitz bestimmter elektrischer bzw. chemischer Eigenschaften dieser Ereignisse das entsprechende Verhalten auslösen, während gewisse relationale Eigenschaften der neuronalen Ereignisse die Strukturierung des Verhaltens bedingen. Dieses Erkenntnis ist die Essenz der Kernthese von Dretskes Aufsatz – nämlich, dass psychologische Erklärungen strukturierende Ursachen und biologische Erklärungen auslösende Ursachen für Verhalten liefern.

Der Hauptunterschied zwischen psychologischen und biologischen Eigenschaften eines zerebralen Ereignisses besteht nach Dretske vornehmlich im extrinsischen bzw. intrinsischen Charakter der jeweiligen Eigenschaften. Während die intrinsischen Eigenschaften die physikalische Beschaffenheit eines Dings repräsentieren, sind die extrinsischen Eigenschaften im Wesentlichen relationale Eigenschaften, d.h. sie sind abhängig von der Beziehung des Trägers der entsprechenden Eigenschaft zu anderen Dingen. Zur Illustration sei nochmals an die intrinsischen und extrinsischen Eigenschaften eines Geldstücks erinnert: so fallen beispielsweise Größe, Dichte, Gewicht, etc. in die Kategorie der intrinsischen Merkmale, wohingegen etwa der Geldwert ein extrinsisches Attribut der Münze darstellt. Obwohl die Unterscheidung zwischen intrinsischen und extrinsischen Eigenschaften die Verschiedenartigkeit von Mentalem und Physikalischem nicht vollständig erschöpft, so ist sie, Dretske zufolge, dennoch ein zentrales Charakteristikum zur Distinktion zwischen beiden.⁴¹⁸ Aus diesem Grund wäre auch ein großer Schritt in Richtung Demonstration der kausalen Relevanz des Mentalen gegangen, wenn es gelänge zu zeigen: „(...) that although it may be a brain's intrinsic properties that are instrumental in triggering bodily behaviour, it is its extrinsic properties that are relevant to structuring this same behaviour. If nothing else, this will at least show that there is nothing about the extrinsicness of the mental that disqualifies it from playing a causal role in the explanation of behaviour.“⁴¹⁹ Dretskes Intention ist es also zu zeigen, dass selbst falls die intrinsischen Eigenschaften eines Gehirnzustandes für das Auslösen von Körperbewegungen verantwortlich sein sollten, es dennoch möglich sein müsste, dass dessen extrinsische Eigenschaften dasselbe Verhalten strukturieren. Dies würde gleichwohl bedeuten, dass der extrinsische Charakter des Mentalen es immerhin nicht davon abschirmt eine kausale Rolle in der Erklärung von Verhalten zu spielen.

⁴¹⁸ Vgl. Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 128ff.

⁴¹⁹ Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 131.

Damit hat Dretske den Grundstein gelegt, um eine Theorie mentaler Verursachung vorzustellen, die seiner Meinung nach kohärent ist. Den Ausgangspunkt seiner Überlegungen bildet folgendes, Verhalten generierendes, System:

I shall describe a system Q whose internal states make it behave in various ways. One particular state – I will label it $R+$ to suggest a positive representational state – is such that tokens of it cause movements of type M . $R+$ is to be understood as an intrinsic state: a state defined by R 's possession of certain intrinsic properties, P_i (that is, $R+ = R$'s having P_i). $R+$ also has certain extrinsic properties, P_e . When Q is a living system, the intrinsic properties are electrical, chemical, and so on. The extrinsic properties of $R+$ can be anything: causal, informational, functional, historical – anything, in fact, that might plausibly be taken to underlie $R+$'s representational content (thereby qualifying Q as inhabiting an appropriate mental state – e.g. believing so-and-so).⁴²⁰

Dretske geht von einem System Q aus, dessen interne Zustände es dazu veranlassen sich auf eine bestimmte Art und Weise zu verhalten. So verursachen beispielsweise die einzelnen Vorkommnisse eines bestimmten positiven repräsentationalen Zustands $R+$ immer Körperbewegungen der Gattung M . Dabei ist $R+$ als ein intrinsischer Zustand zu verstehen – d.h. der Wesenskern von $R+$ ist durch R 's Besitz gewisser intrinsischer Eigenschaften P_i gekennzeichnet. Allerdings weist $R+$ auch bestimmte extrinsische Merkmale P_e auf. Wenn Q ein Lebewesen ist, dann sind die intrinsischen Eigenschaften chemischer, elektrischer,... Natur und seine extrinsischen Charakteristika können ihrem Wesen nach kausal, informationell, funktional, historisch, usw. sein. Der entscheidende Punkt für die Zugehörigkeit zu den extrinsischen Eigenschaften ist, dass sie dem repräsentationalen Gehalt von $R+$ zu Grunde liegen und es Q somit ermöglichen sich in einem mentalen Zustand (wie z.B. zu glauben, dass so und so) zu befinden.

Um zu zeigen, dass die extrinsischen Eigenschaften P_e von $R+$ eine (zumindest teilweise) kausale Erklärung einer Bewegung des Typs M liefern können, schmückt Dretske sein Beispiel vom System Q so aus:

Suppose that system Q is wired so that $R+$ is tokened when, and only when, an object of type O appears nearby. That is, R registers O 's presence by going into its $+$ state – an internal sign or indicator that an O is nearby. O 's presence causing a token of $R+$ in Q can be thought of as Q sensing (perceiving) O . Suppose, furthermore, that O s are dangerous, and being dangerous, Q learns to avoid them. Learning to avoid them is a developmental process in which evasive movements, M , come to be produced when O s are nearby. The only way such movements can be co-ordinated with the presence of an O (so that Q engages in these movements when, but – generally speaking – only when, an O is nearby) is if an internal sign or indicator of O – in this case $R+$ – is *itself* made into a cause of those evasive manoeuvres. The way to co-ordinate M with O is to make O , via $R+$, a cause of M . So, as a result of learning, $R+$, the internal sign of O , becomes a cause (a triggering cause) of M .

Imagine that you know none of this, but that you observe Q (in the wild, as it were) avoiding O . At least you observe Q exhibiting movements M , and you want to know why. Why is Q doing M ?

There are two possible answers – one corresponding to the triggering cause of this behaviour, the other corresponding to its structuring cause.⁴²¹

⁴²⁰ Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 131f.

⁴²¹ Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 132.

Angenommen das System Q ist ein Lebewesen und die einzelnen Vorkommnisse des Zustands R+ treten immer genau dann (und nur dann) ein, wenn ein Objekt des Typs O in der Nähe von Q auftaucht. Die Anwesenheit von Os verursacht also einen R+ Zustand – sozusagen als eine Art innerer Repräsentation der Präsenz eines O – in Q. Diesen Vorgang könnte man mit ‚Q nimmt ein O in seiner unmittelbaren Umgebung wahr‘ paraphrasieren. Geht man zudem davon aus, dass Os für Q gefährlich sind (z.B. weil sie deren Fressfeinde sind), so wird Q dies mit der Zeit begreifen und schließlich lernen Os zu meiden. Dieser Entwicklungsprozess hat zur Folge, dass Q sobald ein O in seine Sichtweite gelangt, Fluchtbewegungen M ausführt. Die einzige Möglichkeit eine konstante Verbindung zwischen den beiden Ereignistypen ‚Fluchtmanöver‘ und ‚Anwesenheit eines O‘ zu schaffen, besteht darin die innere Repräsentation der Präsenz eines O in Q – also R+ – selbst zur Ursache von M zu machen. Der Lernprozess, den Q durchläuft, hat also zur Folge, dass R+ als innerer Platzhalter für die Gegenwart eines O, zu einer auslösenden Ursache von M wird.

Jemand, der über keinerlei Vorwissen verfügt und in freier Natur beobachtet wie Q mit M auf ein O reagiert, könnte sich nun die Frage stellen warum dies so ist. Allerdings gibt es zwei gleichwertige – jedoch fundamental verschiedene – Kausalerklärungen für dieses Phänomen; je nachdem, ob man die auslösende Ursache oder die strukturierende Ursache zur Erklärung des Fluchtverhaltens M heranzieht. Hier zunächst die Darlegung mittels der auslösenden Ursache:

Triggering Causal Explanation: Q is running because it saw (perceived, sensed) an O – because an O approached (and Q saw it). This is a triggering causal explanation because one is merely citing a link in a triggering causal chain (...). The approach of O caused, in Q, an event (we are calling it R+) which, in turn, produced M. Knowing nothing of the intrinsic nature (i.e. the neurobiological character) of this internal triggering event, an observer is forced to specify the triggering cause of the behaviour M by ‘backing up’, so to speak, to the cause of the triggering cause (which is itself, of course, a triggering cause). The approach of O made Q run by causing in Q an event (R+) which triggered movement M. We normally compress this into the simple: Q ran because it saw an O approaching.⁴²²

Die Erklärung mithilfe der auslösenden Ursache begründet das Weglaufen Qs damit, dass es ein O in seiner Umgebung wahrgenommen hat. Genauer gesagt, hat das Herannahen eines Os in Q ein Ereignis verursacht – und zwar R+, eine innere Repräsentation des Umstands, dass ein O in der Nähe ist. Dieses Ereignis (R+) hat wiederum Qs Fluchtbewegungen M hervorgebracht. Da ein Beobachter in der Regel nichts über die genaue intrinsische Beschaffenheit von R+ weiß, sprich, er nicht mit den neurobiologischen Eigenschaften von R+ vertraut ist, muss er zur Erklärung des Verhaltens M gewissermaßen auf die auslösende Ursache der auslösen-

⁴²² Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 133.

den Ursache rekurren. Das Nahen eines *O*s hat *Q* zur Flucht veranlasst, indem es in *Q* das Ereignis *R*⁺ verursacht hat, welches seinerseits die Bewegungen *M* erzeugt hat. Umgangssprachlich wird dieser Sachverhalt mit ‚*Q* ist geflüchtet, weil sich ein *O* näherte‘ wiedergegeben.

Trotz dieser Kausalerklärung ist es möglich, dass es weiterhin nicht vollständig verständlich ist, warum *Q* eigentlich die Flucht ergriffen hat. Vielleicht ist beispielsweise nicht bekannt, dass *O*s gefährlich für *Q* sind. Bzw. selbst falls Kenntnis hinsichtlich dieses Sachverhalts besteht, bleibt weiterhin offen, ob sich *Q* auch dessen bewusst war. Die Art und Weise wie *O* in *Q* repräsentiert ist, ist also von essentieller Bedeutung. In anderen Worten ist nicht so sehr entscheidend *was* *Q* wahrnimmt, sondern vielmehr *welches* (vermeintliche) *Wissen* *Q* hinsichtlich seiner Perzeptionen hat. Die ausschlaggebende Fragestellung in diesem Kontext ist also, warum *Q* beim Anblick eines *O* das Weite gesucht hat. So ist es durchaus vorstellbar, dass zwei unterschiedliche Lebewesen *Q*₁ und *Q*₂ beim Erscheinen eines *O* fliehen. Die Gründe für dieses Fluchtverhalten mögen jedoch in der Tat grundverschieden sein. Während *Q*₁ in *O* z.B. einen Feind erkennt und deshalb den Rückzug antritt, könnte *Q*₂ *O* etwa als Freund erkennen, dem es seit langem noch einen Gefallen schuldet und daher meiden möchte. Nach außen hin löst also das Auftauchen von *O* bei *Q*₁ und *Q*₂ genau das gleiche Verhalten aus – sie entfernen sich fluchtartig – die Gründe für dieses Gebaren können jedoch diametral entgegengesetzt sein. Anders ausgedrückt ist die auslösende Ursache in beiden Fällen identisch, wohingegen die strukturierende Ursache in beiden Konstellationen stark divergiert.⁴²³ Auf das ursprüngliche Beispiel bezogen, beschreibt Dretske die Kausalerklärung mittels der strukturierenden Ursache folgendermaßen:

Structuring Causal Explanation: In supplying the animal's reasons for running from the (external) triggering cause of its flight we are, I submit, providing structuring causes of flight. (...) [T]he observer of *Q* may (after watching *Q* long enough) know that what is triggering *Q*'s behaviour is the sight of *O*. What he doesn't know is why the sight of *O* (the internal *R*⁺ which registers the presence of *O*) causes *M*: evasive movements.

A structuring cause is needed: some causal explanation for why seeing *O* makes *Q* run. What is it that configured *Q* so as to make *R*⁺ (the internal perceptual registration of *O*) a cause of movements of type *M*? Before learning occurred the approach of *O* caused *R*⁺ in *Q*. Before learning to run, before learning that they were dangerous, *Q* could see *O*s approaching. It could see predators before it learned to recognize them as predators. But, at this time, *R*⁺ did not cause *M*. Now it does. Now *Q* runs from the *O*s it sees. Why?

The question about *Q*'s reasons is, I submit, a question about the structuring cause of *Q*'s behaviour, and in the little scenario I have constructed, its answer obviously has to do with *Q*'s learning history. What explains why, during learning, *R*⁺ was recruited as, made into, a cause of *M* is the fact that *R*⁺ was a sign of *O* and the organism had a need to co-ordinate behaviour – in this case evasive movements *M* – with the presence of *O*. *R*⁺ was somehow (and I leave the neurobiological details of learning – what we know of them – to experts) made into a cause of *M* because, by signalling the presence of *O*, *R*⁺ ‘told’ the animal what it needed to know (viz. when to produce *M*) to escape harm. Hence, the internal sign of *O*, *R*⁺ was made into a cause of *M* – solving, at one stroke, the co-ordination problem. If, in order to survive, you have to do *M* when *O*, then you have to make some

⁴²³ Vgl. Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 133.

internal sign of *O* into a cause of *M*. There is no way learning can occur unless it happens, unless *R*⁺, an internal indicator of *O*, is recruited as a cause of *M*. That, indeed, is why animals need internal representations.⁴²⁴

Das Erhellen des Grundes für das Fluchtverhalten von *Q* ist nach Dretske nichts anderes als das Einbringen einer strukturierenden Ursache dafür. Beobachtet ein Feldforscher *Q* in freier Wildbahn, so kann er zwar feststellen, dass das Erblicken von *Os* zu einer Fluchtreaktion führt, ihm ist jedoch nicht (zwangsläufig) klar, warum der Anblick eines *Os* diese Wirkung zur Folge hat. Die Kausalerklärung mithilfe einer strukturierenden Ursache beantwortet die Frage nach dem Warum für *Qs* Fluchtverhalten beim Anblick eines *Os*. Die Kernfrage nach der strukturierenden Ursache von *M* lautet also: Was hat das System *Q* dergestalt eingerichtet, dass *R*⁺, die interne Repräsentation der Wahrnehmung eines *Os*, Bewegungen der Art *M* verursacht? Die Antwort auf diese Frage hängt maßgeblich mit einem Lernprozess zusammen, den *Q* im Laufe der Zeit durchläuft. Dieser Vorgang ist von elementarer Bedeutung für das Überleben von *Q*. Da *Os* Raubtiere sind, in deren Beuteschema *Q* passt, ist es für *Q* von enormer Wichtigkeit die Anwesenheit eines *Os* mit einer unmittelbaren Fluchtreaktion zu beantworten. Eine sehr elegante Methode die Präsenz eines *Os* mit bestimmten Fluchtbewegungen *M* zu verknüpfen, besteht darin, die innere Repräsentation der Gegenwart eines *Os* in *Q* – also *R*⁺ – zur Ursache von *M* zu machen. Der Umstand, dass *R*⁺ als Ursache von *M* etabliert wird, ist unabdingbar für *Qs* Lernprozess. Wird *R*⁺ nämlich nicht zur Ursache von *M* gemacht, dann ist im Gegenzug auch kein Lernen möglich, so Dretske.

Der entscheidende Punkt, der mentale Verursachung nach Dretske überhaupt erst ermöglicht, ist in Folgendem zu sehen:

The important point to notice is that when behaviour is the result of learning of this sort, the structuring cause of the behaviour is *R*'s possession of an extrinsic property (or, if you prefer, the fact that *R*⁺ has this extrinsic property). It is the fact that tokens of *R* (whatever, intrinsically, they may be like) are signs of *O*, indicators of *O*, carriers-of-information about *O*, that explains why a change was wrought in *Q* so as to make (future tokens of) *R*⁺ causes of *M*. Since being a sign of *O*, an indicator of *O*, a carrier-of-information about *O*, are all extrinsic to the internal tokenings of *R*⁺ that trigger *M*, the structuring cause of *Q*'s behaviour is an extrinsic fact about *Q*'s internal states. It may seem (recall our example of the terrorist) as though one and the same state – namely, *R*⁺ – is both triggering and structuring cause of *M*, but this is mistaken. It is *R*'s having the intrinsic, neurobiological properties defining *R*⁺ that is the triggering cause of *M*; it is the fact that *R*⁺ indicates *O*, an extrinsic fact about *R*⁺, that explains why *Q* was reconfigured to make *R*⁺ into a cause of *M*. *R*'s having *P*_{*i*} is the triggering cause of *M*; *R*'s having *P*_{*e*} is the structuring cause.⁴²⁵

Falls also dem Verhalten von *Q* tatsächlich ein solcher Lernprozess zu Grunde liegt, dann ist die strukturierende Ursache der Körperbewegung *M* nichts anderes als *R*'s Besitz einer gewissen extrinsischen Eigenschaft. Denn die für das Lernen entscheidenden Eigenschaften von

⁴²⁴ Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 133f.

⁴²⁵ Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 134f.

R sind extrinsischer Natur. Unabhängig davon welche intrinsischen Merkmale R genau besitzt, ist der Schlüsselfaktor für den Lernvorgang von Q in dem Verhältnis zu seiner Umwelt, genauer gesagt zu Os, zu sehen – also in einer extrinsischen Eigenschaft. Schließlich ist für den evolutionären Prozess, in dem R+ zur Ursache von M wird, allein die Tatsache, dass R in einer gewissen Verbindung zu Os steht, entscheidend. Die kausale Verknüpfung zwischen dem Fluchtverhalten M und den einzelnen Vorkommnissen von R beruht folglich auf dem repräsentationalen Gehalt von R. Damit ist auch die strukturierende Ursache von Qs Verhalten M in einer extrinsischen Eigenschaft von R zu lokalisieren.

Zunächst mag es so scheinen als wären sowohl die auslösende als auch die strukturierende Ursache des Fluchtverhaltens M in ein und demselben Zustand – namentlich R+ – zu suchen. Allerdings zeigt eine eingehende Analyse, dass es R+'s Besitz gewisser intrinsischer (i.e. neurobiologischer) Eigenschaften P_i ist, der die auslösende Ursache von M verkörpert. Wohingegen die Tatsache, dass R+ für Q die Anwesenheit eines O repräsentiert – also (eine Teilmenge) R+'s extrinsische(r) Eigenschafte(n) P_e – die strukturierende Ursache von M darstellt. Anders ausgedrückt, erklärt der repräsentationale Gehalt von R+, warum die einzelnen Vorkommnisse (Token) von R zur Ursache von M wurden.

Am Ende seines Aufsatzes bilanziert Dretske die Vorzüge seiner Theorie mentaler Verursachung im Hinblick auf das Externalismus-Problems wie folgt:

If we embed *Q* in a radically different environment, a world or habitat in which internal tokenings of *R*+, although the same in every intrinsic respect, do not indicate the presence of an *O* (i.e. they have different extrinsic properties), then even if they still cause *M*, the structuring cause (and, hence, the explanation) of this behaviour will be quite different. Perhaps in this different world, tokens of *R*+ cause *M* because they indicate (or indicated) the presence, not of *O*, but of twin-*O*, a much different sort of predator or, perhaps, not a predator at all but just a creature that *Q*, for reasons of its own, has a need (or desire) to avoid. In this possible habitat, *Q*, given time enough and luck, learns to avoid twin-*O*. The causal explanation of *Q*'s behaviour in this novel environment, though it is (in one respect) identical to the behaviour of *Q* in the normal habitat (in both places *Q* avoids or flees from *O*-looking creatures) is quite different. In one case *Q* runs because of its past interaction with *Os* (because his internal *R* indicates the approach of an *O*); in the other case *Q* runs because of its past interaction with twin-*Os* (because, in this different environment, *Q*'s internal *R* indicates the approach of a twin-*O*). On certain theories of representational content (including my own) *Q*'s internal states have different representational content in these two worlds. The behaviours, described in one way, are the same, but they are none the less causally explained by different content. Meanings don't have to be in the head to do their job – even when the job they are doing is the same.⁴²⁶

Um die Effektivität seiner Theorie mentaler Verursachung hinsichtlich des Externalismus-Problems zu demonstrieren, lädt Dretske zu folgendem Gedankenexperiment ein: Angenommen Q findet sich in einer Umgebung wieder, die von der aktuellen Welt insofern divergiert, als dass dort die einzelnen inneren Vorkommnisse von R+, obwohl sie in all ihren intrinsischen Eigenschaften identisch mit den aktual-weltlichen sind, dennoch unterschiedliche ex-

⁴²⁶ Dretske, Fred: „*Mental events as Structuring Causes of Behaviour*“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993, S. 135f.

trinsische Eigenschaften besitzen. Denn die Token von R^+ repräsentieren in dieser möglichen Welt nicht die Anwesenheit eines Os. Vielmehr indizieren sie die Präsenz eines Zwillings-Os – das könnte beispielsweise eine andere Raubtierart als O sein oder sogar ein anderes Lebewesen, das Q aus bestimmten Gründen meiden will bzw. muss. Sofern Q in dieser neuen Umgebung lange genug überlebt, wird es lernen Zwillings-Os aus dem Weg zu gehen. Obwohl nun Qs Verhalten in der aktuellen und in der beschriebenen möglichen Welt in einer Hinsicht identisch ist – in beiden Fällen ergreift Q nämlich die Flucht, sobald es ein O-ähnliches Geschöpf in seiner Umgebung wahrnimmt – ist die kausale Erklärung von Qs Verhalten dennoch unterschiedlich, so Dretske. Im aktual-weltlichen Szenario nimmt Q aufgrund seiner bisherigen Begegnungen mit Os Reißaus, während in der beschriebenen möglichen Welt die Ursache für Qs Fluchtverhalten hingegen in seinen Interaktionen mit Zwillings-Os liegt, die bis dato stattgefunden haben. Der essentielle Unterschied zwischen beiden Gegebenheiten findet sich also in dem Inhalt, den die internen Rs jeweils repräsentieren – sprich, in einem Fall die Anwesenheit eines Os, im anderen die Präsenz eines Zwillings-Os. Gemäß externalistischer Theorien haben die jeweiligen Zustände R von Q in den beiden Welten *unterschiedliche* repräsentationale Gehalte, obwohl sie in all ihren intrinsischen Eigenschaften identisch sind. Dies hat zur Folge, dass Qs Verhaltensweisen in beiden Szenarien, wenngleich sie nach einer bestimmten Art der Beschreibung gleich sind, dennoch mittels unterschiedlicher Gehalte kausal erklärt werden. Diese Schwierigkeit wiederum wird von Dretskes Theorie mentaler Verursachung umgangen. Im Bezug auf Putnams Diktum, Bedeutungen seien nicht im Kopf, konstatiert Dretske, dass sie auch nicht im Kopf sein müssen, um ihre Arbeit zu erledigen, selbst wenn es dieselbe Arbeit ist, die sie zu verrichten haben.

In seiner Gesamtheit stellt Dretskes Ansatz die momentan elaborierteste Antwort auf die in dieser Arbeit vorgestellten und diskutierten Probleme mentaler Kausalität dar. Da Dretske nicht von der Existenz eigenständiger mentaler Substanzen ausgeht, umschiffte seine Theorie das *Interaktionsproblem* in all seinen Facetten. Was das *Problem um die Anomalie des Mentalen* anbelangt, so folgt Dretske Davidson und vertritt genau wie dieser eine token identity theory. Dieser Auffassung zufolge ist die Welt ihrem innersten Wesen nach physikalisch und dementsprechend muss jedes einzelne Vorkommnis eines mentalen Ereignisses mit einem Vorkommnis eines physikalischen Ereignisses identisch sein. Das Kernproblem für derartige Ansätze besteht jedoch darin den Nachweis zu erbringen wie Mentales qua Mentalem kausal wirksam sein kann. Anders ausgedrückt gilt es nachzuweisen, dass mentale Eigenschaften kausale Relevanz besitzen. Dies ist auch die Quintessenz der beiden letzten, in diesem Rahmen

behandelten, Problemkomplexe mentaler Kausalität. Diese sind einerseits das *Exklusionsproblems* und andererseits das *Externalismusproblem*. Im Hinblick auf diese beiden Aspekte der Frage nach der kausalen Relevanz mentaler Eigenschaften kann die Dual-Explanandum Strategie ihre wahre Trumpfkarte ausspielen. Indem Dretske nämlich zwischen der Verursachung von Körperbewegungen und der Verursachung von Verhalten differenziert, kann er damit auch den vermeintlichen Wettbewerb der beiden Ursachen auflösen. Zudem gelingt es Dretske zu zeigen wie extrinsische Eigenschaften einen Beitrag in Kausalerklärungen leisten können.

Literaturverzeichnis

Monographien

- Beckermann, Ansgar: *Analytische Einführung in die Philosophie des Geistes*, 2. überarb. Auflage, Berlin, de Gruyter, 2001.
- Dardis, Anthony: *Mental Causation – The Mind-Body Problem*, New York, Columbia University Press, 2008.
- Descartes, René: *Meditationes de Prima Philosophia / Meditationen über die Erste Philosophie*, Lat. Dt. (Gerhart Schmidt), Stuttgart, Philipp Reclam jun., 1986.
- Foster, John: *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*, London, Routledge, 1991.
- Harbecke, Jens: *Mental Causation – Investigating the Mind's Powers in a Natural World*, Frankfurt, Ontos Verlag, 2008.
- Kim, Jaegwon: *Mind in a Physical World – An Essay on the Mind-Body Problem and Mental Causation*, Cambridge (Mass.), MIT Press, 2000.
- Kim, Jaegwon: *Physicalism, or something near enough*, Princeton, Princeton University Press, 2005.
- Lauth, Bernhard: *Descartes im Rückspiegel: Der Leib-Seele-Dualismus und das naturwissenschaftliche Weltbild*, Paderborn, Mentis, 2006.
- McGinn, Colin: *The Character of Mind*, Oxford, Oxford University Press, 1982.
- Psillos, Stathis: *Causation & Explanation*, Durham, Acumen, 2002.
- Putnam, Hilary: *Philosophical Papers Volume 2: Mind, Language and Reality*, Cambridge, Cambridge University Press, 1975.
- Strawson, Peter Frederick: *Einzelding und logisches Subjekt (Individuals)*, Stuttgart, Philipp Reclam jun., 1972.
- Unger, Peter K.: *All the Power in the World*, Oxford, Oxford University Press, 2005.
- Walter, Sven: *Mentale Verursachung – Eine Einführung*, Paderborn, Mentis, 2006.

Anthologien

- Corcoran, Kevin (Hg.): *Soul, Body, and Survival: Essays on the Metaphysics of Human Persons*, Ithaca, Cornell University Press, 2001.
- Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001.
- Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000.
- Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993.
- Russell, Bertrand: *Mysticism and Logic and other Essays*, 10. Auflage, London, Allen & Unwin, 1951.

Aufsätze

- Block, Ned: „Do Causal Powers Drain Away?“ in: *Philosophy and Phenomenological Research*, Vol. 67, 2003.
- Burge, Tyler: „Individualism and the Mental“ in: *Midwest Studies in Philosophy*, Vol. 4, 1979.
- Crisp, Thomas & Warfield, Ted: „Kim's Master Argument“, in: *Noûs*, Vol. 35:2, 2001.
- Davidson, Donald: ‚Mental Events‘ in: Davidson, Donald: *Essays on Actions and Events*, 2. Auflage, Oxford, Clarendon Press, 2001.
- Dretske, Fred: „Mental events as Structuring Causes of Behaviour“ in: Heil, John & Mele, Alfred (Hrsg.): *Mental Causation*, Oxford, Clarendon Press, 1993.
- Dretske, Fred: „Minds, Machines, and Money: What Really Explains Behavior“ in: Dretske, Fred: *Perception, Knowledge and Belief – Selected Essays*, Cambridge, Cambridge University Press, 2000.
- Dretske, Fred: „Reasons and Causes“ in: *Philosophical Perspectives*, Vol. 3, 1989.
- Fodor, Jerry: „A Modal Argument for Narrow Content“ in: *The Journal of Philosophy*, Vol. 88, 1991.
- Fodor, Jerry: ‚Making Mind Matter More‘ in: *Philosophical Topics*, Vol. 17, Nr. 1, 1989.
- Honderich, Ted: ‚The Argument for Anomalous Monism‘ in: *Analysis*, Vol. 42, 1982.
- Horgan, Terence: „Mental Quausation“ in: *Philosophical Perspectives*, Vol. 3, 1989.

- Jackson, Frank & Pettit, Philip: „*Program Explanation – A General Perspective*“, in: *Analysis*, Vol. 50, 1990.
- Jehle, David: „*Kim against Dualism*“ in: *Philosophical Studies*, Volume 130, 2006.
- Kistler, Max: „*Causation in contemporary analytical philosophy*“, in: *Quaestio*, Vol.2, 2002.
- Le Pore, Ernest & Loewer, Barry: „*Mind Matters*“ in: *The Journal of Philosophy*, Vol. 84, 1987.
- Lewis, David: „*Causation*“ in: *Journal of Philosophy*, Vol. 70 (17), 1973.
- Malcolm, N.: „*The Conceivability of Mechanism*“ in: *Philosophical Review*, Vol. 77, 1968.
- Marras, Ausonio: „*Kim’s Supervenience Argument and Nonreductive Physicalism*“ in: *Erkenntnis*, Vol. 66, 2007.
- O’Connor, Timothy: „*Causality, Mind, and Free Will*“ in: Corcoran, Kevin (Hg.): *Soul, Body, and Survival: Essays on the Metaphysics of Human Persons*, Ithaca, Cornell University Press, 2001.
- Owens, Joseph: „*Content, Causation, and Psychophysical Supervenience*“ in: *Philosophy of Science*, Vol. 60, 1993.
- Richardson, R. C.: „*The ‘Scandal’ of Cartesian Interactionism*“ in: *Mind*, Volume 91, 1982.
- Robb, David: „*The Properties of Mental Causation*“ in: *The Philosophical Quarterly*, Vol. 47, No. 187, 1997.
- Russell, Bertrand: „*On the Notion of Cause*“ in: Russell, Bertrand: *Mysticism and Logic and other Essays*, 10. Auflage, London, Allen & Unwin, 1951.
- Smart, J.J.C.: „*Sensations and Brain Processes*“ in: *Philosophical Review*, Vol. 68, 1959.
- Sosa, Ernest: „*Mind-Body Interaction and Supervenient Causation*“ in: *Midwest Studies in Philosophy*, Vol. 9, 1984.
- Yablo, Stephen: „*Mental Causation*“ in: *The Philosophical Review*, Vol. 101, 1992.

Internetquellen

- Stanford Encyclopedia of Philosophy (<http://plato.stanford.edu/>)
 - o insbesondere: Robb, David & Heil, John.: Eintrag ‚Mental Causation‘ (<http://plato.stanford.edu/entries/mental-causation/>)
- Internet Encyclopedia of Philosophy (<http://www.iep.utm.edu/>)
 - o insbesondere: Yoo, Julie: Eintrag ‚Mental Causation‘ (<http://www.iep.utm.edu/mental-c/>)